

UNIVERZA V LJUBLJANI  
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Anže Javornik

**Ocenjevanje zanesljivosti posameznih napovedi  
inkrementalnih modelov**

DIPLOMSKO DELO

UNIVERZITETNI ŠTUDIJSKI PROGRAM RAČUNALNIŠTVO IN  
INFORMATIKA

Ljubljana, 2016



UNIVERZA V LJUBLJANI  
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Anže Javornik

**Ocenjevanje zanesljivosti posameznih napovedi  
inkrementalnih modelov**

DIPLOMSKO DELO

UNIVERZITETNI ŠTUDIJSKI PROGRAM RAČUNALNIŠTVO IN  
INFORMATIKA

MENTOR: prof. dr. Igor Kononenko  
Ljubljana, 2016



To delo je ponujeno pod licenco *Creative Commons Priznanje avtorstva-Deljenje pod enakimi pogoji 2.5 Slovenija* (ali novejšo različico). To pomeni, da se tako besedilo, slike, grafi in druge sestavine dela kot tudi rezultati diplomskega dela lahko prosto distribuirajo, reproducirajo, uporabljajo, priobčujejo javnosti in predelujejo, pod pogojem, da se jasno in vidno navede avtorja in naslov tega dela in da se v primeru spremembe, preoblikovanja ali uporabe tega dela v svojem delu, lahko distribuira predelava le pod licenco, ki je enaka tej. Podrobnosti licence so dostopne na spletni strani [creativecommons.si](http://creativecommons.si) ali na Inštitutu za intelektualno lastnino, Streliška 1, 1000 Ljubljana.



Izvorna koda diplomskega dela, njeni rezultati in v ta namen razvita programska oprema je ponujena pod licenco *GNU General Public License*, različica 3 (ali novejša). To pomeni, da se lahko prosto distribuira in/ali predeluje pod njenimi pogoji. Podrobnosti licence so dostopne na spletni strani <http://www.gnu.org/licenses>.



Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Ocenjevanje zanesljivosti posameznih napovedi inkrementalnih modelov

Tematika naloge:

Pri uporabi naučenih modelov za napovedovanje je pomembno tudi ocenjevanje zanesljivosti posameznih napovedi. Za statične modele so ocene zanesljivosti že dodobra razvite in se v praksi tudi uporabljajo. Področje ocenjevanja zanesljivosti inkrementalnih modelov pa je še v razvoju. V okviru diplomske naloge razvijte metode za ocenjevanje zanesljivosti inkrementalnih modelov in jih testirajte na napovedih različnih inkrementalnih algoritmov strojnega učenja z uporabo več umetnih in realnih podatkovnih množic. Pri tem analizirajte časovno in prostorsko zahtevnost razvitih algoritmov in jih primerjajte z metodami za ocenjevanje zanesljivosti klasičnih (neinkrementalnih) modelov. Analizirajte tudi, kako se kvaliteta ocen zanesljivosti spreminja skozi čas.





## IZJAVA O AVTORSTVU DIPLOMSKEGA DELA

Spodaj podpisani Anže Javornik sem avtor diplomskega dela z naslovom:

*Ocenjevanje zanesljivosti posameznih napovedi inkrementalnih modelov*  
(angl. *Reliability estimation of individual predictions for incremental models*)

S svojim podpisom zagotavljam, da:

- sem diplomsko delo izdelal samostojno pod mentorstvom prof. dr. Igorja Kononenka
- so elektronska oblika diplomskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko diplomskega dela,
- soglašam z javno objavo elektronske oblike diplomskega dela na svetovnem spletu preko univerzitetnega spletnega arhiva.

V Ljubljani, dne 25. maja 2016

Podpis avtorja:



## **Zahvala**

Posebna zahvala je namenjena Igorju Kononenku, ki me je kljub večkratnim razočaranjem znova in znova sprejel pod svoje okrilje kot mentor. Pri delu mi je pomagal z nasveti in potrpežljivimi razlagami. Darko Pevec mi je pomagal pri načrtu dela, izbiri programskih knjižnic in pri preverjanju pravilnosti. Zoranu Bosniću se zahvaljujem za vse napotke pri statistični obdelavi rezultatov. Za konec pa še zahvala tudi vsem najdražjim, za vso pomoč in podporo.



*Svoji dragi Petrci.*



# Kazalo

<b>Povzetek.....</b>	<b>1</b>
<b>Abstract .....</b>	<b>2</b>
<b>1 Uvod .....</b>	<b>3</b>
1.1 Motivacija in cilji .....	3
1.2 Pregled področja .....	4
1.3 Pregled vsebine.....	4
<b>2 Inkrementalno strojno učenje in klasifikacija .....</b>	<b>5</b>
2.1 Inkrementalni naivni Bayesov klasifikator.....	5
2.2 Inkrementalna odločitvena drevesa .....	6
2.3 Inkrementalni bagging in boosting odločitvenih modelov .....	7
<b>3 Ocenjevanje zanesljivosti.....</b>	<b>9</b>
3.1 Strategija prilagajanja inkrementalnih ocen .....	11
3.2 Inkrementalna ocena lokalne napake.....	15
3.3 Metoda lokalnega prečnega preverjanja .....	16
3.4 Inkrementalna varianca modela bagging .....	18
<b>4 Metodologija testiranja ocen zanesljivosti .....</b>	<b>21</b>
4.1 Statistične metode.....	21
4.2 Testiranje na fiksnih domenah.....	22
4.3 Testiranje na generatorjih .....	22
<b>5 Rezultati testiranj .....</b>	<b>23</b>
5.1 Poskusi na fiksnih domenah .....	23
5.2 Poskusi na generatorjih brez zamika .....	24
5.3 Poskusi na generatorjih z zamikom .....	27
<b>6 Aplikacija na realni domeni .....</b>	<b>31</b>
6.1 Lastnosti domene.....	31
6.2 Testiranje ocen zanesljivosti.....	32

<b>7</b>	<b>Diskusija in zaključki.....</b>	<b>35</b>
7.1	Pregled rezultatov .....	35
7.2	Nadaljnje delo .....	36
	<b>Literatura.....</b>	<b>37</b>
	<b>Dodatek A .....</b>	<b>39</b>



## Seznam uporabljenih kratic

kratica	Angleško	slovensko
<b>CA</b>	classification accuracy	klasifikacijska točnost
<b>BAG</b>	bootstrap aggregating	združevanje klasifikatorjev
<b>BAGV</b>	Variance of bagging model	varianca modela bagging
<b>CNK</b>	Evaluation of local fault	ocena lokalne napake
<b>kNN</b>	nearest neighbour method	metoda najbližjih sosedov
<b>LCV</b>	local cross validation method	metoda lokalnega prečnega preverjanja
<b>NB</b>	naive Bayes	naivni Bayesov klasifikator
<b>DT</b>	decision tree	odločitveno drevo kot klasifikator
<b>MOA</b>	massive online analysis	masivna sprotna analiza
<b>iNB</b>	incremental naive Bayes	inkrementalni naivni Bayesov klasifikator
<b>HT</b>	Hoeffding Tree	klasifikator Hoeffding Tree
<b>OBAG</b>	Oza Bag	klasifikator Oza Bag
<b>OBOOST</b>	Oza Boost	klasifikator Oza Boost



# Povzetek

**Naslov:** Ocenjevanje zanesljivosti posameznih napovedi inkrementalnih modelov

Diplomsko delo obravnava ocenjevanje zanesljivosti posameznih napovedi inkrementalnih modelov. Namen ocene zanesljivosti je obogatiti napoved modela z dodatno informacijo. Ta dodatna informacija ima lahko kritični pomen, zlasti če imajo napačne napovedi lahko hude posledice.

Klasični modeli napovedovanja so zgrajeni na primerih problematike, za katero podajajo svojo napoved. Klasične metode ocenjevanja zanesljivosti se gradijo na podoben način na primerih, ki predstavljajo učno množico modelu napovedovanja. Na tak način zgrajeni modeli in ocene zanesljivosti so na kratek rok uspešni. V daljšem obdobju uporabe pa najverjetneje pride do sprememb pravil v problematiki, ki negativno vplivajo na uspešnost klasičnih napovednih modelov in ocen zanesljivosti. V realnosti je večina problematik takih, da se pravila v njih spreminjajo. Grajenje novih klasičnih modelov za prilagajanje problematiki je zamudno ali sploh ne pride v poštev, če je podatkovni tok tak, da lahko primer vidimo samo enkrat – npr. mrežna komunikacija. Za analizo takih problemskih domen se zato odločimo za uporabo inkrementalnih modelov napovedovanja, ki se znajo prilagajati spremembam pravil problematike.

V diplomskem delu predstavimo tri znane metode ocenjevanja zanesljivosti in predlagamo njihove različice z inkrementalnim značajem. Znane metode in razvite različice smo preizkusili na dvaindvajsetih fiksnih množicah in osmih podatkovnih tokovih (generatorjih) brez zamika in z zamikom, ki nam predstavlja spremembo pravil v problemski domeni. Dobljene ocene smo statistično obdelali. Pri statistični analizi nas je zanimala statistična značilnost korelacije dejanske napake z oceno. V tem delu pridobljeni rezultati kažejo, da predlagane inkrementalne metode ocenjevanja zanesljivosti ponujajo skoraj enake rezultate kot klasične metode na fiksnih problemskih domenah in mnogo boljše rezultate pri podatkovnih tokovih, še posebej pri tokovih z zamikom. Najboljše rezultate nudi ocena iCNK. Dobre rezultate nudi tudi ocena iBAGV. Inkrementalna različica ocene LCV, iLCV, pa ne odstopa od svoje klasične različice.

**Ključne besede:** inkrementalno strojno učenje, klasifikacija, inkrementalno ocenjevanje zanesljivosti, klasifikacijska točnost

# Abstract

**Title:** Reliability estimation of individual predictions for incremental models

The thesis deals with the reliability estimates of individual predictions of incremental models. The purpose of reliability estimates is to enrich the models predictions with additional information. This information might have a critical meaning, especially if wrong predictions can have serious consequences.

Classical prediction models are built on examples from a problem domain for which they provide predictions. Classical methods of the reliability estimation are constructed in a similar way, on examples representing the learning set for the prediction model. Models and reliability estimate scores built in this way in short term provide good results. In the long term use there is a high probability that a change happens in the problem domain that adversely affects the performance of conventional predictive models and estimates of reliability. Building a new predictive model to adapt to the changes can be time consuming or even not an option, if the data stream is such that we can observe each example only once – e.g. network communication. Therefore, for the analysis of such problem domains we choose to use the incremental prediction models, which are able to adapt to changes.

In this thesis we present three known methods of the reliability estimation and propose their versions with an incremental character. These methods were tested on twenty-two fixed problem domains and eight data streams (generators) with and without drift, which represents a change in the problem domain. The resulting estimates were statistically analysed. In the statistical analysis, we were interested in statistically significant correlations between factual error and the reliability assessment. Obtained results show that the proposed incremental methods of reliability estimation offer almost the same results in fixed problem domains and much better results in data streams, especially if the drift is applied. Best results are offered by the reliability estimate iCNK. Good results are also offered by the reliability estimate iBAGV. An incremental version of the reliability estimate LCV, iLCV, does not deviate from its classical version.

**Keywords:** incremental machine learning, classification, incremental reliability estimates, classification accuracy

# 1 Uvod

Strojno učenje se ukvarja z izgradnjo modelov, ki so namenjeni napovedovanju. Iz nekega predhodnega znanja (učnih primerov) o določeni problematiki se zgradi znanje, ki opisuje problematiko (modelira učne primere).

Proces učenja je prilagajanje modela s ciljem, da nalogo opravlja bolje. Rezultat učenja je znanje, ki je lahko samo množica zapomnjenih podatkov, algoritem ali pa neka množica napotkov (pravil) za bolj učinkovito reševanje nalog. Takemu znanju pogosto rečemo kar model. Za model zahtevamo, da čim bolj ustreza vhodnim podatkom in predznanju [1].

Zgoraj opisani proces učenja predpostavlja, da so vsi podatki na voljo na začetku učenja. Učenju, kjer se znanje spreminja sproti s sprejemanjem novih podatkov, pa pravimo inkrementalno učenje. V algoritmu inkrementalnega učenja se poskuša najti najmanjšo potrebno spremembo znanja, tako, da ta še ustreza vsem do sedaj videnim podatkom.

Model, kot rezultat učenja, se lahko uporabi bodisi za reševanje novih primerov (napovedovanje) bodisi za razumevanje problematike. Napovedovanje je proces, kjer se primeru, ki ni bil vključen v proces učenja, poskusi napovedati iskana vrednost. Glede na naravo vrednosti ločimo probleme na klasifikacijske (diskretna vrednost) in regresijske (zvezna vrednost).

S procesom napovedovanja je tesno povezan pojem zanesljivosti. V primeru, kjer ima napoved lahko resne posledice, bi želeli oceniti tudi, kako zanesljiva je ta napoved. Zato želimo vsako napoved obogatiti z informacijo o njeni točnosti. Ker pa točnost napovedi za neznan primer ne moremo definirati, jo samo ocenjujemo.

## 1.1 Motivacija in cilji

Pogosto se sama problematika s časom spreminja. Lahko se spreminja distribucija v učni množici ali pa pravila v problemski domeni. Primeri takih področij so mrežni promet, podatki senzorjev ali sporočilne storitve. Podatki v takih primerih pogosto hitro postanejo stari in se je temu potrebno prilagoditi. Takim podatkom pravimo podatkovni tokovi. Za primer lahko vzamemo iskanje novic na spletu, kjer se vsak dan pojavi ogromno število novih novic, iskalnik pa mora iskati najnovejše. V takih primerih nam klasični modeli skozi čas dajejo vedno slabše

napovedi. Za take primere uporabimo inkrementalne modele, ki so se sposobni takim spremembam prilagajati.

Medtem ko je področje ocenjevanja zanesljivosti klasičnih modelov že precej raziskano, je področje ocenjevanja zanesljivosti inkrementalnih modelov še dokaj nedotaknjeno. Cilj diplomske naloge je razviti metode za ocenjevanje zanesljivosti inkrementalnih modelov in jih testirati na različnih modelih in bazah podatkov. Prav tako nas zanima, kakšni so zanje časi izračuna zanesljivosti? in kako se ocene zanesljivosti? skozi čas spreminjajo. Razvite metode želimo primerjati z metodami za ocenjevanje zanesljivosti klasičnih modelov.

## 1.2 Pregled področja

Diplomsko delo spada v področje ocenjevanja zanesljivosti napovedi, podrobneje na področje ocenjevanja zanesljivosti napovedi inkrementalnih modelov. Za izračun ocene zanesljivosti lahko uporabimo pristop, kjer upoštevamo model napovedovanja ali pa ne. Pristopi, kjer upoštevamo model napovedovanja, upoštevajo lastnosti modela in so tako manj splošni kot pristopi, kjer je osnova za izračun opazovanje sprememb napovedi ob spremembah učne množice ali atributov. Osnova nekaterih metod za ocenjevanje zanesljivosti je transdukcija. Transduktivno sklepanje je sklepanje od konkretnega h konkretnemu – iz učnih primerov sklepamo o vrednosti novega primera.

Že znane metode za ocenjevanje zanesljivosti napovedi, ki so našteje v [2], niso primerne za uporabo pri inkrementalnih modelih. Osnova teh ocen je učna množica. Ker pa se le ta pri podatkovnih tokovih ne preneha več in vsebinsko spreminja, je izračun obstoječih ocen zanesljivosti velikokrat prepočasen in zato v sistemih, kjer je potreben hiter odgovor, neuporaben. V ta namen želimo nekatere že obstoječe metode za ocenjevanje zanesljivosti prilagoditi inkrementalnim modelom.

## 1.3 Pregled vsebine

V 2. poglavju najprej na kratko navedemo nekaj inkrementalnih klasifikacijskih modelov, ki smo jih uporabili v nadaljnjem delu. Nato v 3. poglavju predstavimo nekaj znanih ocen zanesljivosti in njihove razvite inkrementalne različice. V 4. poglavju te razvite metode in obstoječe metode testiramo in med seboj primerjamo. Za vrednotenje ocen zanesljivosti napovedi uporabimo statistične teste. V 5. poglavju razvite metode poizkusimo na podatkovnih tokovih in ovrednotimo stopnjo korelacije med ocenami zanesljivosti in klasifikacijsko napako. Nato si ogledamo, kako se spreminja stopnja korelacije skozi čas. V zadnjem poglavju sledijo še diskusija, pregled rezultatov in zaključki.

## 2 Inkrementalno strojno učenje in klasifikacija

Pri klasičnem strojnem učenju učni algoritem kot vhod dobi predznanje in množico učnih primerov. Z uporabo predznanja se preišče prostor možnih hipotez in kot rezultat vrne zaključno hipotezo, ki se nato uporabi za napovedovanje vrednosti novih vhodnih podatkov. Napoved je lahko diskretne ali zvezne narave. V primeru, da je ciljna napoved diskretne narave govorimo o klasifikaciji, drugače o regresiji. Inkrementalno strojno učenje se od klasičnega razlikuje po tem, da ima za vhod samo predznanje. Vsak primer pa se nato uporabi za prilagoditev zaključne hipoteze.

Najpogostejše vhodni primer opišemo z atributno predstavitevjo, ki je v [1] definirana z množico atributov  $A$ :

$$A = A_i, A = 0 \dots a$$

Atribut  $A_i$  je lahko diskreten

$$\forall A_i: \mathcal{V}_i = V_1, \dots, V_n$$

ali pa zvezen

$$\forall A_i: \mathcal{V}_i = [Min_i, Max_i]$$

En učni primer lahko zapišemo kot vektor vrednosti atributov, kjer je z  $r_{(j)} = v_{(0,j)}$  označen ciljni razred:

$$u_j = \{r_{(j)}, v_{(1,j)}, \dots, v_{(a,j)}\}$$

Če rešujemo klasifikacijski problem, je razred diskreten atribut, če pa rešujemo regresijski problem, je razred zvezni atribut.

V diplomskem delu uporabljamo že implementirane klasifikatorje v knjižnici MOA [3]. V naslednjih podpoglavjih so klasifikatorji na kratko opisani.

### 2.1 Inkrementalni naivni Bayesov klasifikator

Naivni Bayesov klasifikator naredi napoved pri predpostavki, da so vrednosti različnih atributov pri danem razredu pogojno neodvisne. Formulo klasifikatorja dobimo s pomočjo Bayesovega pravila:

$$P(r_k | V) = P(r_k) \prod_{i=1}^a \frac{P(r_k | v_i)}{P(r_k)}$$

Naloga učnega algoritma je s pomočjo učne množice aproksimirati verjetnosti na desni strani enačbe. Znanje naivnega Bayesovega klasifikatorja je tabela aproksimacij apriornih verjetnosti razredov  $P(r_k)$  in tabela pogojnih verjetnosti  $P(r_k | v_i)$  razredov  $r_k$  pri dani vrednosti  $v_i$  atributa  $A_i$ .

Pri inkrementalnem učenju pa mora algoritem določiti novo apriorno verjetnost in pogojno verjetnost razredov glede na nov testni primer in njegovo klasifikacijo. V osnovi se vsak novi klasificiran primer uporabi za popravo znanja. Prednost tega je, da lahko ohranja predhodno videno znanje.

Trenutno učno množico označimo z  $D$ ,  $T$  je množica novih primerov in obstaja primer

$$x_p = (A_0, \dots, A_a) \in T$$

z apriornimi verjetnostmi za razrede  $P(r_k)$  in pogojnimi verjetnostmi  $P(r_k | v_i)$ . Popravna formula za apriorne vrednosti razredov je tako

$$P'(r_k) = \begin{cases} \frac{s}{s+1} P(r_k) + \frac{1}{1+s} & ; A_0 = r_k \\ \frac{s}{s+1} P(r_k) & ; A_0 \neq r_k \end{cases}$$

kjer

$$s = |D| + |T|$$

in popravna formula za pogoje verjetnosti

$$P(r_k, v_i) = \begin{cases} \frac{m}{1+m} P(r_k, v_i) + \frac{1}{1+m} & ; A_0 = r_k \text{ in } A_i = v_i \\ \frac{m}{1+m} P(r_k, v_i) & ; A_0 = r_k \text{ in } A_i \neq v_i \\ P(r_k, v_i) & ; A_0 \neq r_k \end{cases}$$

kjer

$$m = |A_i| + \text{število}(r_j)$$

kjer  $|A_i|$  predstavlja število vrednosti atributa  $A_i$  in  $\text{število}(r_j)$  predstavlja število primerov, katerih razred je  $r_k$  [4].

V knjižnici MOA najdemo implementacijo inkrementalnega naivnega Bayesovega algoritma. Ker naivni Bayesov algoritem računa pogoje verjetnosti za diskretne spremenljivke, je potrebna predhodna diskretizacija zveznih atributov.

## 2.2 Inkrementalna odločitvena drevesa

Odločitveno drevo je sestavljeno iz notranjih vozlišč, ki ustrezajo atributom, vej, ki ustrezajo podmnožicam vrednosti atributov, in listov, ki ustrezajo razredom. Ena pot v drevesu od korena do lista ustreza enemu odločitvenemu pravilu. Pri tem so pogoji, pari atribut – podmnožica vrednosti, konjunktivno povezani [5].



Naloga učnega algoritma je zgraditi odločitveno drevo iz učne množice primerov in sicer tako, da za vsak nivo najde trenutno najpomembnejši atribut, ki najbolj deli množico primerov, in ga uporabi za koren trenutnega poddrevesa. Inkrementalni učni algoritmi odločitvenih dreves pa popravljajo drevo za vsak nov učni primer. Cilj algoritma za inkrementalno učenje odločitvenega drevesa je, da na enaki množici učnih primerov zgradi enako drevo, kot neinkrementalni algoritem. Primer inkrementalnega učnega algoritma je ID5R, ki za nov primer najde najpomembnejši atribut  $A_i$ , ga dvigne proti korenu drevesa in drevo popravi [6].

V našem delu smo uporabili inkrementalna odločitvena drevesa, ki izkoriščajo Hoeffdingovo neenakost, ki pravi, da so že majhni vzorci pogosto lahko dovolj dobri za izbiro optimalnega odločitvenega atributa. Bolj natančno Hoeffdingova neenakost pravi, da z verjetnostjo  $1 - \delta$  lahko napovemo povprečno vrednost neke spremenljivke  $x$  v obsegu  $R$ , ki se ne bo razlikovala od dejanske povprečne vrednosti pri  $n$  ponovitvah za več kot [7]

$$\varepsilon = \sqrt{\frac{R^2 \ln(\frac{1}{\delta})}{2n}}$$

V knjižnici MOA je implementirano več različic Hoeffdingovih dreves. V diplomskem delu je uporabljena osnovna implementacija, kjer je v listu drevesa za strategijo napovedi uporabljen večinski razred.

## 2.3 Inkrementalni bagging in boosting odločitvenih modelov

Tradicionalni algoritmi za nadzorovano učenje podajo svojo klasifikacijo glede na posamezen osnovni odločitveni model, kot je recimo odločitveno drevo ali naivni Bayes. Skupinski učni algoritmi pa združujejo napovedi mnogih odločitvenih osnovnih modelov, kjer je vsak naučen na tradicionalen način. Bagging in boosting sta dobro poznana skupinska učna algoritma, za katere se je pokazalo, da vplivata na izboljšanje rezultatov v primerjavi z individualnimi modeli [8].

Bagging dela tako, da iz originalne učne množice  $N$  primerov generira  $M$  učnih množic velikosti  $N$  tako, da naključno izbere primere iz originalne učne množice z vračanjem. Vsaka tako zgrajena učna množica vsebuje vsak učni primer  $K$ -krat, kjer

$$P(K = k) = \binom{N}{k} \left(\frac{1}{N}\right)^k \left(1 - \frac{1}{N}\right)^{N-k}$$

kar je binomska distribucija. Z  $N \rightarrow \infty$  se distribucija  $K$  približuje k Poisson(1) distribuciji

$$K \sim \frac{e^{-1}}{k!} \pi r^2$$

To lastnost izkorišča inkrementalni algoritem bagging. Za vsak nov primer, za vsak osnovni model v množici  $M$ , izbere primer  $K \sim \text{Poisson}(1)$  krat in posodobi model  $M$  [8].

Boosting je kompleksnejši proces grajenja množice osnovnih modelov  $M$ . Vsak osnovni model  $h_1, \dots, h_m$  je naučen z uteženo množico učnih primerov. Utež določa klasifikacijska napaka predhodnega osnovnega modela  $h_{m-1}$ . Natančneje, primeri, ki so napačno klasificirani v osnovnem modelu  $h_{m-1}$ , imajo večjo težo v učni množici za naslednji osnovni model  $h_m$ , pri čemer uteži vseh napačno klasificiranih primerov sestavljajo polovico teže učne množice, ostalo polovico pa pravilno klasificirani primeri. Tudi v tem primeru inkrementalna različica učnega algoritma boosting uporablja Poissonovo vzorčenje za približke oteževalnega algoritma. Algoritem deluje na podoben način kot bagging, le da ko osnovni model napačno klasificira primer, se Poissonov distribucijski parameter, vezan na ta primer, poveča za naslednji osnovni model, drugače pa zmanjša [8].

V knjižnici MOA je na voljo implementacija tako bagging-model OzaBag in boosting-model OzaBoost in nekaj njunih različic. V diplomskem sta uporabljeni različici z NB kot osnovnim klasifikatorjem.

### 3 Ocenjevanje zanesljivosti

Rezultat strojnega učenja je znanje, ki se ga lahko uporabi za boljše razumevanje problematike ali za napovedovanje. V praksi pa sama napoved ni vedno dovolj. Velikokrat ima odločitev lahko tudi resne posledice. Zato bi radi vsako napoved tudi podkrepili z njeno točnostjo. Vendar o točnosti težko govorimo, saj imamo opravka s problemom, ki ga ne poznamo. Poznamo pa podobne primere, s katerimi smo se že srečali v preteklosti. Zato lahko le sklepamo o točnosti rešitve tega novega problema. Kako točna ali zanesljiva pa je naša rešitev? S tem vprašanjem se ukvarja področje ocenjevanja zanesljivosti. Takoj nam na misel pride ideja: več ko primerov poznamo, z večjo zanesljivostjo bomo lahko podali novo napoved. Seveda pa je preverjanje celotnega prostora že znanih primerov zelo zamudno in v praksi neuporabno. Zato se največkrat osredotočimo na le najbolj podobne ali bližnje primere. Za mero podobnosti vzamemo razdaljo med primeri.

Klasični modeli napovedovanja se učijo le enkrat na vnaprej znani učni množici primerov. Ta učna množica in tako zgrajeno znanje je statično in se s časom ne spreminja. Pri inkrementalnem učenju pa se vsak nov primer uporabi za posodobitev obstoječega znanja. Ker se znanje spreminja, nam bodo statični algoritmi za ocene zanesljivosti s časom dajali vedno slabše rezultate. Znanje nam bo naraščalo, medtem ko bomo oceno zanesljivosti osnovali na starem znanju. Tako bomo lahko dobili zelo točne napovedi ampak z zelo slabo oceno zanesljivosti. Če želimo spodobne ocene zanesljivosti, moramo te algoritme prilagoditi inkrementalnemu učenju. V nadaljnjih podpoglavjih predstavimo nekaj že znanih ocen zanesljivosti in predlagamo njihove inkrementalne različice, ki jih v naslednjih poglavjih testiramo.

Pri razvoju inkrementalnih ocen zanesljivosti moramo zagotoviti delovanje v omejenem prostoru in (skoraj) konstantnem času. To pomeni, da prostor, ki ga algoritem potrebuje, in čas za izračun, ne smeta preseči neke zgornje mejne vrednosti. Tako bo naš največji izziv pri razvoju inkrementalnih ocen zanesljivosti, kako se spopasti z vedno naraščajočim številom primerov, saj število le teh v največji meri vpliva tako na prostor kot na čas izračuna, ne da bi izgubili informacijo, ki jo primer nosi.

Klasifikatorji, glede na svojo naravo, vračajo svoje napovedi ali v obliki verjetnostnih distribucij (v primeru klasifikacijskih problemov), ali v obliki vrednosti (v primeru regresijskih problemov). Tako je rezultat napovedi klasifikacijskega problema

$$P \in \mathbb{P} \leftrightarrow P = (p_1, \dots, p_n) \text{ in } \sum_{i=1}^n p_i = 1; \forall i : p_i \in [0, 1]$$

in rezultat napovedi regresijskega problema

$$P \in \mathbb{P} \leftrightarrow P = (p) \text{ in } p \in \mathbb{R}$$

Oceno zanesljivosti želimo omejiti na interval  $[0, 1]$ , kjer 0 pomeni popolnoma nezanesljivo napoved in 1 popolnoma zanesljivo napoved. Razdaljo med dvema napovedma  $f(P, Q)$ , za katero velja

1. funkcija  $f(P, Q)$  je realna funkcija,
2.  $\forall P, Q \in \mathbb{P}^n : f(P, Q) = 0 \leftrightarrow P = Q$ ,
3. za klasifikacijske probleme  $\forall P, Q \in \mathbb{P}^n : 0 \leq f(P, Q) \leq 1$

lahko merimo na več načinov:

- manhatanska razdalja

$$f_{man}(P, Q) = \sum_i \frac{|p_i - q_i|}{2}$$

- evklidska razdalja

$$f_{evk}(P, Q) = \frac{\sqrt{\sum_i (p_i - q_i)^2}}{\sqrt{2}}$$

- maksimalna razdalja

$$f_{max}(P, Q) = \max_i |p_i - q_i|$$

- kot med porazdelitvama

$$\phi(P, Q) = \cos^{-1} \left( \frac{\sum_i p_i q_i}{\sqrt{\sum_i p_i^2} \sqrt{\sum_i q_i^2}} \right)$$

$$f_{kot}(P, Q) = \frac{\phi(P, Q)}{\frac{\pi}{2}}$$

- kosinus kota med porazdelitvama

$$f_{cos}(P, Q) = 2 \cos^{-1} \left( \frac{\sum_i p_i q_i}{\sqrt{\sum_i p_i^2} \sqrt{\sum_i q_i^2}} \right)$$

- Bhattacharyy razdalja

$$f_{bha}(P, Q) = \sum_i \sqrt{p_i q_i}$$

- skalarni produkt

$$f_{ska}(P, Q) = \sum_i p_i q_i$$

- harmonična srednja vrednost

$$f_{har}(P, Q) = \sum_i \frac{2p_i q_i}{p_i + q_i}$$

- Hellingerjeve razdalja

$$f_{hel}(P, Q) = \sum_i \frac{(\sqrt{p_i} - \sqrt{q_i})^2}{2}$$

- separacijska razdalja

$$f_{sep}(P, Q) = \max_i \left(1 - \frac{p_i}{q_i}\right)$$

- Jensen – Shannon divergenca

$$f_{jsd}(P, Q) = \sum_i \frac{p_i \log\left(\frac{2p_i}{p_i + q_i}\right) + q_i \log\left(\frac{2q_i}{p_i + q_i}\right)}{2}$$

Ker e razdalja komplementarna funkcija naši željeni oceni zanesljivosti, našo oceno zanesljivosti definiramo kot  $1 - \text{razdalja}$ . Pri regresijskih problemih je za skladnost potrebno napovedi predhodno normalizirati.

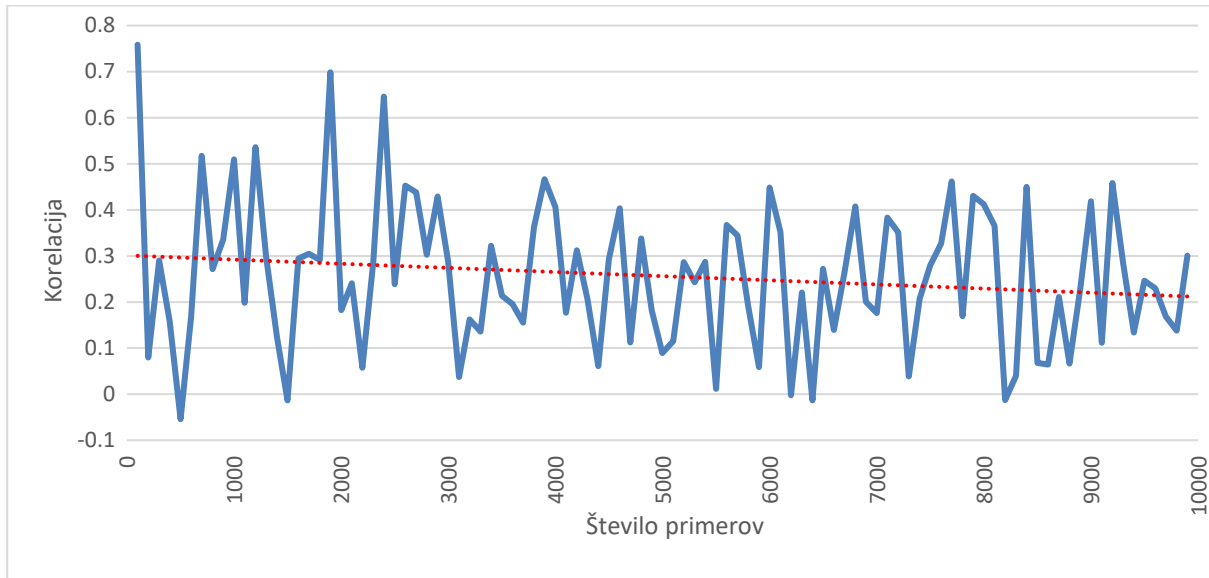
### 3.1 Strategija prilagajanja inkrementalnih ocen

Klasične ocene zanesljivosti temeljijo na nekem predznanju, ki se ga določi enkrat in potem uporablja za izračun zanesljivosti napovedi. Pri grajenju take ocene zanesljivosti imamo dva problema:

- Katere primere vključiti v predznanje?
- Kaj narediti, če se pravila v problematiki spremenijo?

Prvi problem nam neposredno vpliva na kvaliteto ocene in hitrost njenega izračuna. Več primerov, vključenih v predznanje, lahko pomeni boljšo oceno vendar počasnejši izračun in večjo zahtevo po prostoru. Drugi problem pa vpliva na kvaliteto ocene v prihodnosti. V primeru

sprememb v problematiki nam primeri, vključeni v predznanje ocene, ne predstavljajo več trenutnega stanja problematike in tako povzročajo napačne izračune. Ta problem je prikazan na sliki 3.1. Iz slike lahko razberemo, da kvaliteta ocene pada.



Slika 3.1: Spreminjanje korelacije med dejansko napako napovedi in oceno zanesljivosti CNK(5) na vsakih 100 primerov generatorja SEA, kjer si je ocena CNK(5) zapomnila prvih 100 primerov.

Za reševanje obeh zgornjih problemov je potrebno vpeljati strategijo prilagajanja inkrementalnih ocen. Strategija prilagajanja mora zadostovati dvema pogojema:

- Mora biti prostorsko omejena
- Mora imeti skoraj konstantni čas izračuna

Najpreprostejša strategija prilagajanja je drseče okno. Pri tej strategiji si nove primere zapomnimo in stare pozabljamo. Pomanjkljivost te strategije je, da starega znanja ne ohranja. Taka strategija je uporabljena pri inkrementalni oceni BAGV, ki je opisana v razdelku 3.4.

Sedaj pa definirajmo še kompleksnejšo strategijo prilagajanja, ki ohranja predznanje. Najprej je potrebno poudariti, da predznanje sedaj niso več primeri  $(x_i, C_i)$ , ampak opisi  $(x'_i, C'_i)$ . Naj bo

$$\mathbb{P} = \{(x'_1, C'_1), \dots, (x'_k, C'_k)\}$$

množica opisov, kjer je  $x'_a$  nek opis primerov, katerih označba je  $C'_a$ . Za vsak nov primer  $(x, -)$  s klasifikacijo  $K$  se moramo odločiti, ali ga združimo v že obstoječ opis, ali pa bomo z njim začeli generirati nov opis. Odločitev o tem bomo osnovali na razdalji  $f(x, x'_a)$  in analizo prostorske porazdelite  $\mathbb{P}$ .

Definirajmo množico vseh primerov  $P$ , kjer je  $C$  razred prevladujoče napovedi

$$\Omega^C = \{P_i\}; P_i \in \mathbb{P} \text{ in } \max_{index} P_i.C = C$$

Definirajmo najdaljšo razdaljo v prostoru primerov z istim prevladujočim klasifikacijskem razredom

$$\omega^C = \max_{i,j} f(P_i, P_j) = \max_{i,j} f(P_i.x', P_j.x'); P_i, P_j \in \Omega^C$$

in minimalno razdaljo med novim primerom s klasifikacijo  $(x, K)$  in prostorom

$$\omega_x = \min_i f(x, P_i); P_i \in \Omega^C \text{ in } C = \max_{index} K$$

in naj bo

$$P_{min} = \{x', C'\} \in \mathbb{P}$$

za katerega velja, da je  $\omega_x$  minimalna. Novi primer  $(x, -)$  s klasifikacijo  $K$  zapišemo kot  $(x, K)$ .

Dodajanje novega primera potem definiramo kot

$$(x, K) + \mathbb{P} = \begin{cases} \mathbb{P} \cup \{x, K\}, & \frac{\omega_x}{\omega^C} > \tau \\ P_{min} = P_{min} + \{x, K\}, & \frac{\omega_x}{\omega^C} \leq \tau \end{cases}$$

kjer je  $(P_{min} + \{x, K\})$  združevanje dveh opisov, ki ga definiramo kot

$$P_i + P_j = \{P_i.x' + \lambda * (P_i.x' - P_j.x'), P_i.C' + \lambda * (P_i.C' - P_j.C')\}$$

Parameter  $\lambda$  nam določa hitrost prilagajanja in  $\tau$  nam določa koeficient dodajanja.

Da ne bi prekoračili maksimalnega števila opisov, moramo, v primeru dodajanja novega vpisa v množico opisov še poskrbeti za njeno krčenje. V primeru, da v množici  $\mathbb{P}$  prekoračimo število opisov  $N$ , najdemo dva med seboj najbližja si opisa

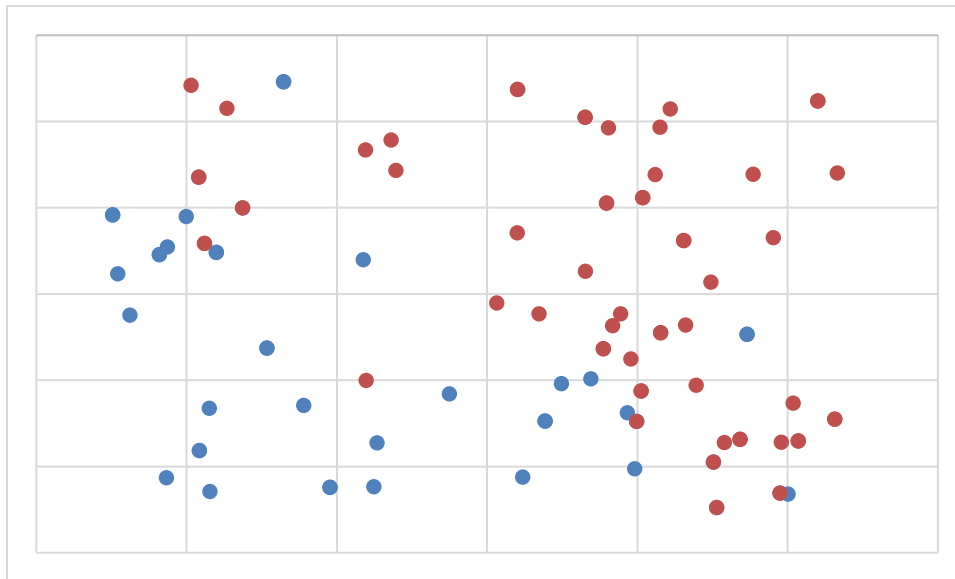
$$\min_{i,j,K} f(P_i, P_j); P_i, P_j \in \Omega^K$$

Krčenje množice primerov definiramo kot

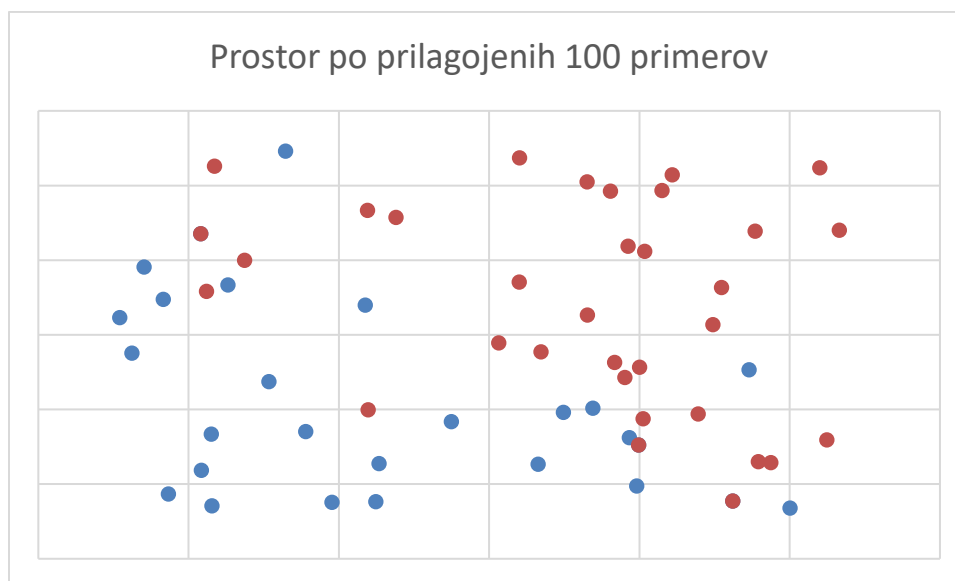
$$\mathbb{P} = f(x) = \begin{cases} \mathbb{P}, & |\mathbb{P}| \leq N \\ (\mathbb{P} \setminus \{P_i, P_j\}) \cup (P_i + P_j), & |\mathbb{P}| > N \end{cases}$$

kjer je  $(P_i + P_j)$  združevanje dveh opisov.

Sliki 3.2 in 3.3 prikazujeta razliko med navadnim pomnjenjem in pomnjenjem z uporabo zgoraj definirane strategije prilagajanja.



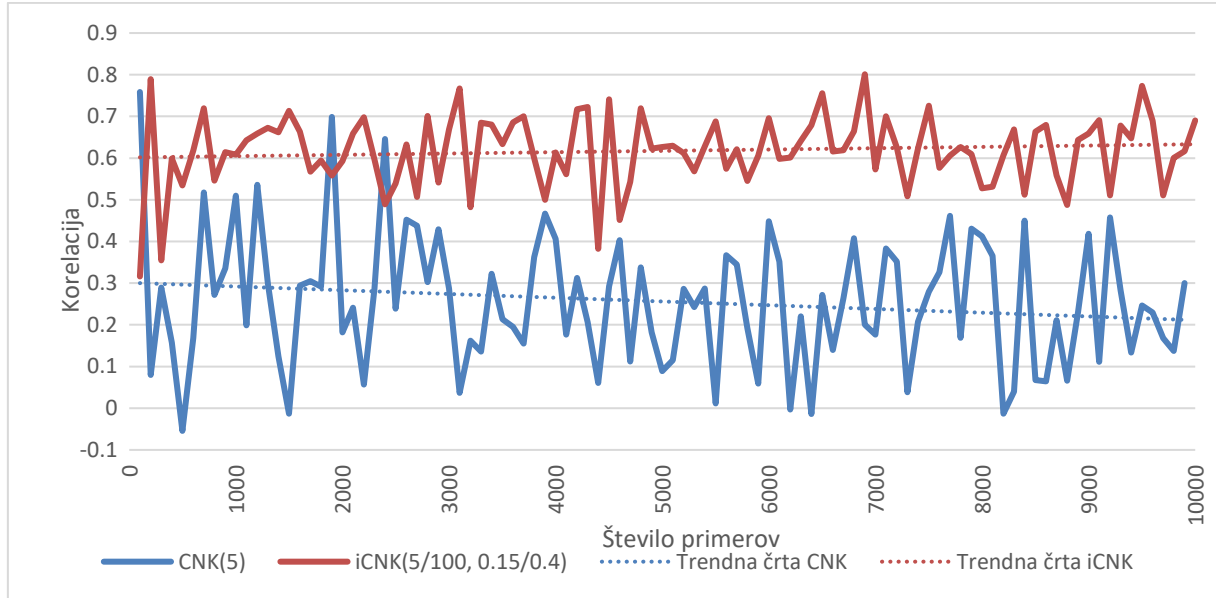
Slika 3.2: Porazdelitev primerov v prostoru po prvih 100 zapomnjenih primerih generatorja SEA.



Slika 3.3: Porazdelitev primerov v prostoru po prvih 100 primerov dodanih s strategijo prilagajanja pri  $\tau = 0.15$  in  $\lambda = 0.4$  (prostor vsebuje 59 opisov).

To strategijo prilagajanja smo uporabili pri inkrementalnih različicah ocen CNK in LCV. Zgoraj predstavljen problem prilagajanja ocene smo tako obrnili in izboljšali kakovost ocene, kot je prikazano v sliki 3.4 .





Slika 3.4: Spreminjanje korelacije med dejansko napako napovedi in ocenama zanesljivosti CNK(5) ter iCNK(5/100, 0.15/0.4) na vsakih 100 primerov generatorja SEA, kjer si je ocena CNK(5) zapomnila prvih 100 primerov.

### 3.2 Inkrementalna ocena lokalne napake

Klasična ocena lokalne napake je definirana kot sledi. Naj bo  $K$  klasifikatorjeva napoved za dani neoznačen primer  $(x, -)$  in  $N = \{(x_1, C_1), \dots, (x_k, C_k)\}$ , kjer je  $C_i$  pravilna označba  $i$ -tega najbližjega sosedu. Ocena lokalne napake za neoznačeni testni primer je definirana kot povprečna mera bližine med  $k$  najbližjimi sosedi in klasifikatorjevo napovedjo:

$$CNK = \frac{\sum_{i=1}^k f(C_i, K)}{k}$$

Pri razvoju algoritma inkrementalne ocene lokalne napake uporabimo v razdelku 3.1 opisano strategijo prilagajanja, kjer imamo množico

$$\mathbb{P} = \{(x'_1, C'_1), \dots, (x'_k, C'_k)\}$$

kjer je  $x'_a$  nek opis primerov, katerih označba je  $C'_a$ .

Definiramo inkrementalno oceno lokalne napake kot povprečno mero bližine med  $k$  najbližjimi opisi in klasifikatorjevo napovedjo, ki jo računa Algoritem 3.1:

$$iCNK(\lambda, \tau, N) = \frac{\sum_{i=1}^k f(C_i^k, K)}{k}$$

Prostorska zahtevnost podanega algoritma je omejena navzgor, saj velja

$$|\mathbb{P}| \leq N$$

kjer je  $N$  maksimalno število opisov. Prav tako je časovna zahtevnost podanega algoritma konstanta, saj ni nikoli potrebno preveriti več kot  $|\mathbb{P}|$  opisov pri izračunu ocene in nikoli več kot  $|\mathbb{P}|^2$  opisov pri dodajanju novega opisa.

**Algoritem 3.1 iCNK****Definicije:**  $N$  – maksimalno število opisov $\tau$  – koeficient dodajanja $\lambda$  – koeficient prilagajanja $\mathbb{P}$  – množica opisov $k$  – število najbližjih sosedov**Vhod:** primer s klasifikacijo  $(x, K)$ **Izhod:** ocena zanesljivosti

1. najdi množico  $k$  najbližjih sosedov  $\mathbb{P}_k$
2. izračunaj oceno  $iCNK(\lambda, N) = \frac{\sum_{i=1}^k f(C_i, K)}{k}$
3. posodobi  $\mathbb{P}$  z  $(x, K)$
4. if  $|\mathbb{P}| > N$ 
  5. najdi primera  $P_i, P_j \in \mathbb{P}$ , za katera velja  $\min_{i,j} f(P_i, P_j)$  in  $P_i.C = P_j.C$
  6. v  $\mathbb{P}$  nadomesti primera  $P_i, P_j$  z  $P_i + P_j$
7. end if
8. vrni vrednost izračunane ocene

**3.3 Metoda lokalnega prečnega preverjanja**

$LCV$  temelji na postopku prečnega preverjanja [2]. Izračun za neoznačen primer  $(x, -)$  je sledeč:

1. določimo množico najbližjih sosedov  $N = \{(x_1, C_1), \dots, (x_k, C_k)\}$
2. za vsakega soseda  $(x_i, C_i)$ 
  1. naredimo klasifikacijski model  $M$  na množici  $N \setminus (x_i, C_i)$
  2. za  $x_i$  z modelom  $M$  izračunamo napoved  $K_i$
  3. za  $x_i$  izračunamo razdaljo  $f(C_i, K_i)$
3. izračunamo oceno zanesljivosti

$$LCV = \sum_{i=1}^k \frac{f(C_i, K_i)}{k}$$

Naš inkrementalni algoritem za izračun zanesljivosti po metodi prečnega preverjanja bo tudi deloval s strategijo prilagajanja, kjer imamo množico opisov

$$\mathbb{P} = \{(x'_1, C'_1), \dots, (x'_k, C'_k)\}$$

kjer  $x'_i$  in  $C'_i$  predstavljata opis primera in njeno klasifikacijsko razporeditev.

Algoritem inkrementalne ocene zanesljivosti po metodi prečnega preverjanja *iLCV* je podan kot Algoritem 3.2.

---

**Algoritem 3.2** *iLCV*


---

**Definicije:**  $N$  – število opisov na klasifikacijski razred

$\tau$  – koeficient dodajanja

$\lambda$  – koeficient prilagajanja

$\mathbb{P}$  – množica opisov

$k$  – število najbližjih sosedov

$\mathcal{M}$  – osnovni inkrementalni model

**Vhod:** primer s klasifikacijo  $(x, K)$

**Izhod:** ocena zanesljivosti

1. najdi množico  $k$  najbližjih sosedov  $\mathbb{P}_k$
  2. for  $(x_i, C_i)$  in  $\mathbb{P}_k$
  3.       zgeneriraj model  $M_i = \text{new}(\mathcal{M})$  in ga nauči  $\mathbb{P}_k \setminus (x_i, C_i)$
  4.       z modelom  $M_i$  izračunaj napoved  $K_i$  za  $x_i$
  5.       izračunaj razdaljo  $f(C_i, K_i)$
  6.     end for
  7.     izračunaj oceno  $iLCV = \sum_{i=1}^k \frac{f(C_i, K_i)}{k}$
  8.     posodobi  $\mathbb{P}$  z  $(x, K)$
  9.     if  $|\mathbb{P}^K| > N$
  10.       najdi primera  $P_i, P_j \in \mathbb{P}$ , za katera velja  $\min_{i,j} f(P_i, P_j)$  in  $P_i.C = P_j.C$
  11.       v  $\mathbb{P}$  nadomesti primera  $P_i, P_j$  z  $P_i + P_j$
  12.     end if
  13.     vrni vrednost izračunane ocene
-

### 3.4 Inkrementalna varianca modela bagging

Agregiran klasifikator, ki je sestavljen iz  $m$  klasifikacijskih modelov, podaja svojo napoved kot povprečje posameznih napovedi

$$K = \sum_{i=1}^m \frac{K_i}{m}$$

Ocena zanesljivosti BAGV pa je definirana kot varianca klasifikacijskih distribucij

$$BAGV = \frac{1}{m} \sum_{i=1}^m (K_i - K)^2$$

Inkrementalno različico BAGV ocene zanesljivosti definiramo na prostoru

$$\mathcal{M} = \{M_1, \dots, M_n, \dots, M_{n+\max(1, \eta * n)}\}$$

kjer je  $M_i$  inkrementalni model,  $\eta$  je koeficient pozabljanja in  $n$  predhodno določeno število modelov. Za strategijo prilagajanja uporabimo malenkostno prirejeno drseče okno. Nov neoznačen primer  $(x, -)$  s klasifikacijo  $K$  uporabimo za inkrementalno učenje  $(1 + \kappa)$  naključno izbranih modelov iz podmnožice  $\mathcal{M}^{learn}, \mathcal{M}^{learn} \subsetneq \mathcal{M}$  z vračanjem. Pri prvem učenju so v množici  $\mathcal{M}^{learn}$  elementi  $\mathcal{M}$  z indeksom 1 do  $m$  v vseh nadaljnjih pa preostali. Prav tako uvedemo pozabljanje na način, da bomo prvih  $m * \eta$  modelov nadomestili z novimi na vsakih  $\frac{m}{\eta}$  novih primerov. Na vrednost ocene zanesljivosti pa bo vplivalo samo prvih  $n$  modelov. To si lahko predstavljamo kot podaljšan seznam  $m + m * \eta$  modelov, ki ima navidezno mejo pri  $m$ -tem elementu, ki deli seznam na levo in desno stran. Dokler ni leva stran dovolj naučena, nove primere uporabimo za učenje modelov na levi strani – to je inicializacija ocene. Nato pa vso učenje delamo samo še na desni strani. Ko je desna stran pripravljena, vzamemo enako število modelov, kot jih je na desni strani, iz začetka leve strani, jih nadomestimo z novimi in postavimo na rep desne strani. Algoritem je podan kot Algoritem 3.3.

---

**Algoritem 3.3** iBAGV

---

**Definicije:**  $\mathcal{M}$  – množica klasifikacijskih modelov $\kappa$  – koeficient prilagajanja $\eta$  – koeficient pozabljanja $n$  – število modelov

bagInit – ali je bag modelov že zgrajen

count – števec za prilagajanje množice  $\mathcal{M}$ **Vhod:** primer s klasifikacijo  $(x, K)$ **Izhod:** ocena zanesljivosti

1. določi množico uporabnih modelov  $\mathcal{M}^u = \{M_i\}$ , kjer je  $M_i \in \mathcal{M}$  in  $i \leq n$
  2. for  $(M_i)$  in  $\mathcal{M}^u$
  3.       izračunaj klasifikacijo  $K_i$  z  $M_i$  za  $x$
  4. end for
  5. izračunaj oceno  $iBAGV = \frac{1}{m} \sum_{i=1}^m (K_i - K)^2$
  6. definiraj množico modelov  $\mathcal{M}^{learn} = \begin{cases} \mathcal{M}^u, & \text{not bagInit} \\ \mathcal{M}/\mathcal{M}^u, & \text{bagInit} \end{cases}$ , kateri so primerni za učenje
  7. naključno izberi  $\mathcal{M}^\kappa = |\mathcal{M}^{learn}| * (1 + \kappa)$  modelov z vračanjem iz  $\mathcal{M}^{learn}$
  8.  $\mathcal{M}^\kappa$  modele nauči primera  $(x, K)$
  9. povečaj števec count za 1
  10. if  $count \geq n$
  11.       prvih  $n * \eta$  modelov iz  $\mathcal{M}$  nadomesti z novimi in jih prestavi na konec  $\mathcal{M}$
  12.       nastavi count = 0 in bagInit = True
  13. end if
  14. vrni vrednost izračunane ocene
-



## 4 Metodologija testiranja ocen zanesljivosti

Ocene zanesljivosti smo testirali na 14 standardnih testnih domenah iz obsežnega repozitorija UCI Machine Learning Repository [6], 8 generatorjih, implementiranih v MOA z in brez zamika ter z 4 inkrementalnimi klasifikacijskimi modeli implementiranimi v MOA, opisanimi v 2. poglavju. Pregled lastnosti domen in generatorjev je predstavljen v tabelah A.1 in A.2 v dodatku A. Inkrementalne ocene smo primerjali s klasičnimi različicami, in sicer:

- CNK z uporabo najbližjih 5 sosedov – CNK(5)
- BAGV s 50 modeli – BAGV(50)
- LCV z uporabo najbližjih 5 sosedov – LCV(5)

S testiranjem nekaj različnih vrednosti parametrov inkrementalnih ocen na generatorjih in inkrementalnim Bayesovim klasifikatorjem smo poiskali tiste, ki v povprečju ponujajo najboljše korelacije med dejansko napako in oceno zanesljivosti. Vrednosti parametrov, ki imajo enak pomen kot v klasičnih različicah, smo obdržali enake. Podrobni rezultati testov so predstavljeni v tabelah A.3 – A.8 v dodatku A. Z dejansko napako so najboljše korelirale spodnje različice inkrementalnih ocen:

- Inkrementalni CNK z  $k = 5$ ,  $N = 100$ ,  $\lambda = 0,65$  in  $\tau = 0,2$  – iCNK(5/100, 0,2/0,65)
- Inkrementalni BAGV z  $N = 50$ ,  $\kappa = 0,3$  in  $\eta = 0,3$  – iBAGV(50, 0,3/0,3)
- Inkrementalni LCV z  $k = 5$ ,  $N = 100$ ,  $\lambda = 0,7$  in  $\tau = 0,25$  – iLCV(5/100, 0,25/0,7)

Nato smo te pridobljene različice ocen primerjali s klasičnimi ocenami zanesljivosti na fiksnih domenah in na domenah z zamikom. V primeru generatorjev smo simulirali fiksno domeno tako, da smo vzeli prvih 200 primerov.

### 4.1 Statistične metode

Dobljene izračune ocen zanesljivosti ovrednotimo z uporabo statističnih metod. Izračunane ocene zanesljivosti primerjamo z dejansko napako napovedi tako, da med njima izračunamo Pearsonove korelacijske koeficiente, katere podkrepimo z statistično značilnostjo.

#### 4.1.1 Pearsonovi korelacijski koeficienti

S korelacijskem koeficientom (korelacija) merimo odvisnost dveh spremenljivk. Pearsonove korelacijske koeficiente računamo na podlagi kovariance in standardnih odklonov vrednosti opazovanih spremenljivk.

#### 4.1.2 Statistična značilnost korelacijskega koeficienta

Statistična značilnost korelacijskega koeficienta nam pove kako (ne)verjetno je, da se dan korelacijski koeficient pojavi, če v populaciji ni nobene povezave. Statistično značilnost izražamo v verjetnostnih nivojih – npr. je značilno pri  $p = 0,05$ . Manjša kot je vrednost  $p$ , bolj značilna je povezava.

#### 4.2 Testiranje na fiksnih domenah

Testiranje na fiksnih domenah smo izvajali na 14 standardnih domenah in 8 generatorjih brez zamika s štirimi inkrementalnimi klasifikacijskimi modeli. Pri generatorjih smo se omejili na prvih 200 primerov. Testiranje ocen zanesljivosti je potekalo v obliki prečnega preverjanja, na način izloči enega – vsakemu izločenemu primeru iz učne množice je bila izračunana napoved, dejanska napaka in ocena zanesljivosti. Na ta način smo dobili izračune za vsak primer iz domene.

Te izračune ocen in dejanskih napak smo statistično ovrednotili.

#### 4.3 Testiranje na generatorjih

Testiranje na generatorjih smo izvajali na 8 generatorjih z in brez zamika s štirimi inkrementalnimi klasifikacijskimi modeli. Pri klasičnih ocenah zanesljivosti smo prvih 100 primerov uporabili za inicializacijo ocene. Testiranje ocen je potekalo na 3000 primerih (2900 v primeru klasičnih ocen). Vsakemu novemu primeru je bila izračunana napoved, dejanska napaka in ocene zanesljivosti.



## 5 Rezultati testiranja

Namen naših testov je ovrednotiti razvite metode ocenjevana zanesljivosti. Na štirinajstih testnih učnih množicah in osmih generatorjih z in brez zamika izračunamo ocene zanesljivosti vseh napovedi klasifikatorjev iNB, HT, OBAG, OBOOST in zabeležimo dejansko napako napovedi. Izračunane vrednosti, s pomočjo statističnih testov, ovrednotimo.

Najprej si ogledamo test ocen zanesljivosti na fiksnih domenah in jih primerjamo. Nato ocene testiramo še na generatorjih brez in z zamikom.

### 5.1 Poskusi na fiksnih domenah

Testiranje ocen smo opravili na štirinajstih standardnih domenah in osmih generatorjih brez zamika. Pri generatorjih smo uporabili prvih dvesto primerov. Testiranje je potekalo po principu prečnega preverjanja, na način izloči enega. Povzetek rezultatov prikazuje tabela 5.1, celotni rezultati pa so v tabelah A.9 – A.12 v dodatku A.

Tabela 5.1: Odstotek značilno pozitivnih in negativnih korelacij med ocenami zanesljivosti in dejansko napako napovedi.

[%]	CNK	BAGV	LCV	iCNK	iBAGV	iLCV
	+/-	+/-	+/-	+/-	+/-	+/-
iNN	95/0	73/0	45/0	86/0	68/0	45/5
HT	95/0	73/0	55/0	86/0	68/5	55/5
OBAG	95/0	86/0	55/0	91/0	68/0	50/5
OBOOST	95/0	77/0	32/23	86/0	64/0	23/14
<b>povprečje</b>	<b>95/0</b>	<b>77/0</b>	<b>47/6</b>	<b>87/0</b>	<b>67/1</b>	<b>43/7</b>

Rezultati v tabeli 5.1 kažejo na to, da so inkrementalne različice klasičnih ocen z njimi primerljive. Vidimo namreč, da imajo zelo podobno statistično značilnost. Pričakovano so inkrementalne ocene malenkostno slabše, saj je njihov prostor ne zajema vseh učnih primerov prečnega preverjanja za razliko od klasičnih ocen.

Pri vseh ocenah zanesljivosti in inkrementalnih klasifikacijskih modelov prevladuje število značilno pozitivnih korelacij. Iz rezultatov je tudi razvidno, da ocene dajejo podobne rezultate pri vseh inkrementalnih klasifikacijskih modelih, razen LCV in iLCV, ki imata pri Oza Boost občutno slabše rezultate.

Tabela 5.2: Delež boljše, slabše, enake ali neznačilne statistične značilnosti v primerjavi z neinkrementalnimi različicami ocen

[%]	boljše	slabše	enako	neznačilno
iCNK	8	35	41	16
iBAGV	3	8	52	36
iLCV	1	6	28	65

Iz tabele 5.2 razberemo, da inkrementalne ocene v večini korelirajo enako kot njihove klasične različice. Ocena iCNK ima v 35% poskusov nižji korelacijski koeficient od ocene CNK in v 8% višji. V 41% primerov pa imata oceni iCNK in CNK po z-testu enak korelacijski koeficient in v 16% vsaj eden od koeficientov ni statistično značilen, zato nadaljnja primerjava ni smotrna. Največjo podobnost kažeta oceni BAGV in iBAGV.

## 5.2 Poskusi na generatorjih brez zamika

Testiranje ocen smo opravili na osmih generatorjih brez zamika. Testirali smo na tri tisoč primerih. Pri uporabi klasičnih ocen smo prvih sto primerov uporabili za njihovo inicializacijo. V tabelah A.13 – A.16 v dodatku A so prikazani podrobni rezultati testiranj na generatorjih za različne modele. Spreminjaje korelacijskega koeficienta skozi število primerov je prikazano grafično v slikah A.1 – A.4 v dodatku A.

Tabela 5.3: Odstotek značilno pozitivnih in negativnih korelacij med oceno zanesljivosti in dejansko napako napovedi.

[%]	CNK +/-	BAGV +/-	LCV +/-	iCNK +/-	iBAGV +/-	iLCV +/-
iNN	100/0	100/0	88/0	100/0	88/12	88/0
HT	100/0	88/0	88/0	88/0	75/12	88/0
OBAG	100/0	88/0	100/0	100/0	88/12	88/0
OBOOST	100/0	88/0	75/12	100/0	75/12	75/0
<b>povprečje</b>	<b>100/0</b>	<b>91/0</b>	<b>88/3</b>	<b>97/0</b>	<b>82/12</b>	<b>85/0</b>

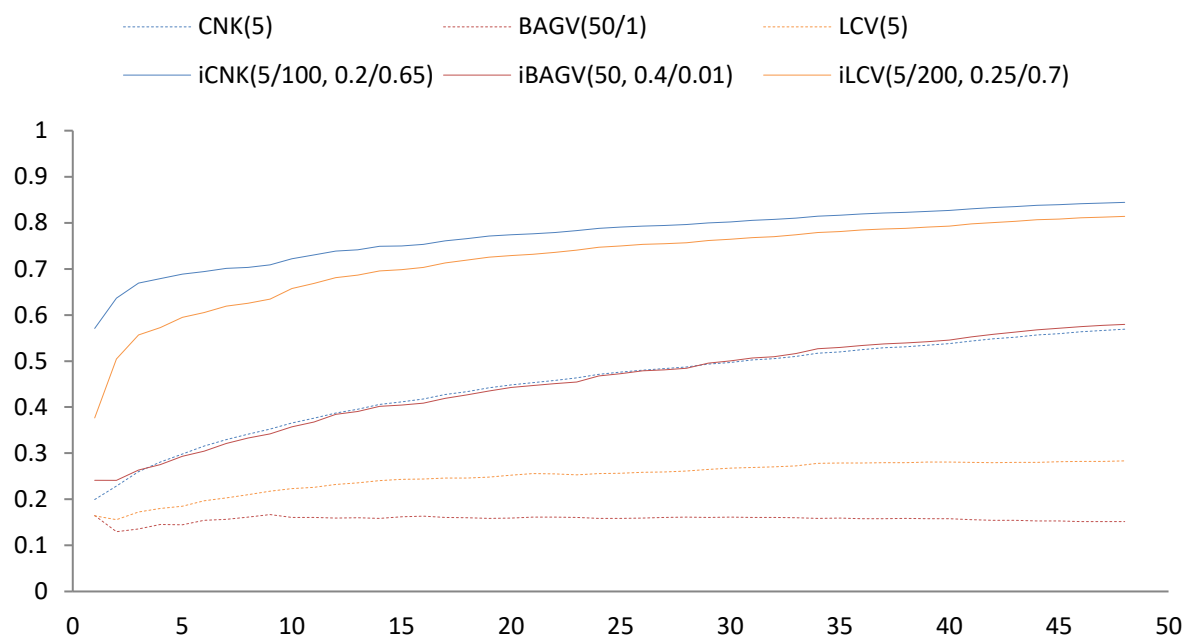
Rezultati v tabeli 5.3 nakazujejo na veliko podobnost med klasičnimi ocenami in njihovimi inkrementalnimi različicami. Vse ocene v večini primerov korelirajo statistično značilno pozitivno. Izstopa ocena iBAGV, ki v 12% korelira značilno negativno. V primerjavi s podatki v tabeli 5.1 lahko vidimo, da so se inkrementalne ocene zelo približale svojim klasičnim različicam, kar nakazuje na njihovo inkrementalno naravo. Visok delež statistične značilnosti ni presenečenje, saj smo izvedli test na 3000 primerih. Pomembnejši rezultati so predstavljeni v tabeli 5.4.

Tabela 5.4: Delež boljše, slabše, enake ali neznačilne statistične značilnosti v primerjavi z neinkrementalnimi različicami ocen

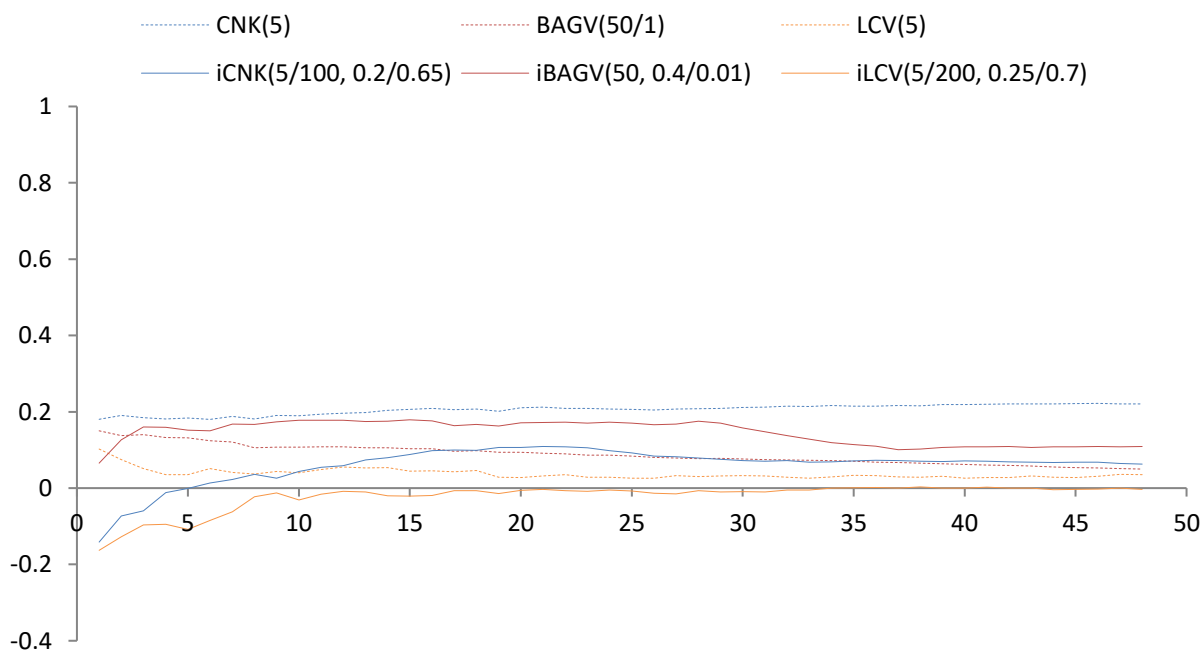
[%]	boljše	slabše	enako	neznačilno
iCNK	72	22	3	3
iBAGV	59	16	12	12
iLCV	19	25	38	19

Tabela 5.4 nam pove, da ima ocena iCNK v primerjavi z CNK v 72% poskusov statistično značilen korelacijski koeficient višji, v 22% nižji, v 3% po z-testu enak in v 3% vsaj eden od koeficientov ni statistično značilen, zato nadaljnja primerjava ni smotrna. Ocena iBAGV ima v 59% višji statistično značilen korelacijski koeficient, v 16% nižji in v 12% po z-testu enak oceni BAGV. Ocen LCV in iLCV pa imata v 38% po z-testu enak korelacijski koeficient, v 25% ima višjega klasična različica in v 19% inkrementalna.

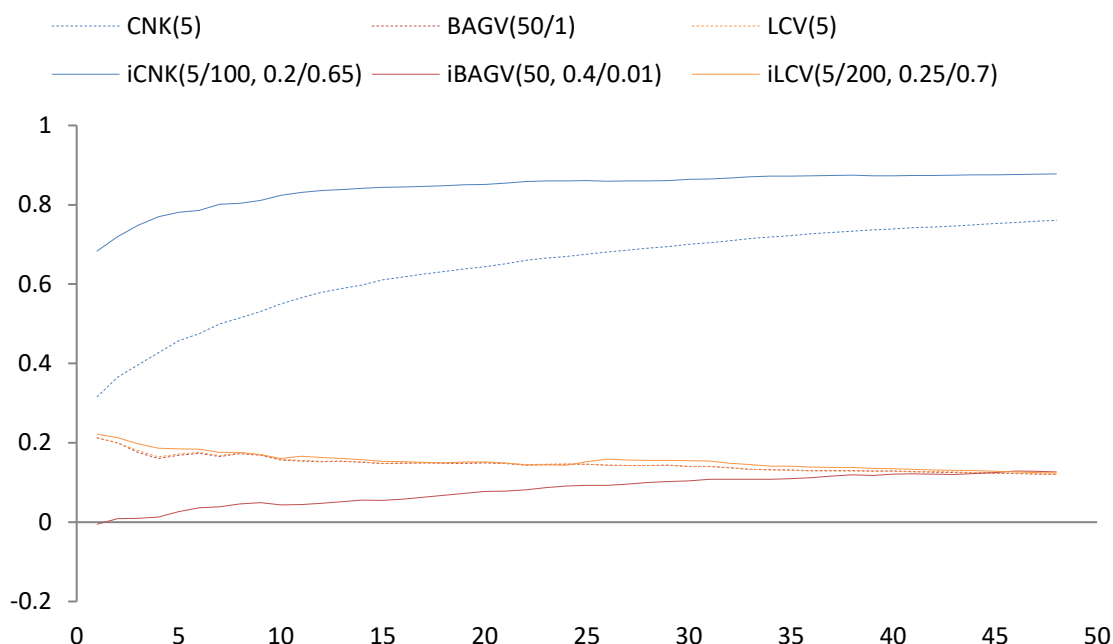
Na slikah 5.1 – 5.3 je prikazanih nekaj primerov gibanja korelacije skozi primere. Vsi primeri so v slikah A.1 – A.4.



Slika 5.1: Spreminjaje korelacijskega koeficienta pri generatorju STAGGER na vsakih 60 primerov pri uporabi klasifikatorja naivni Bayes.



Slika 5.2: Spreminjaje korelacijskega koeficienta pri generatorju Agrawal na vsakih 60 primerov pri uporabi klasifikatorja naivni Bayes.



Slika 5.3: Spreminjaje korelacijskega koeficienta pri generatorju Hyperplane na vsakih 60 primerov pri uporabi klasifikatorja naivni Bayes.

Pridobljeni rezultati nam kažejo na to, da kljub temu da se v problemski domeni pravila ne spreminjajo, nam inkrementalne ocene ponujajo boljši izračun zanesljivosti. Neinkrementalne ocene namreč uporabljajo za svoj izračun samo primere, na katerih so bile zgrajene, kar pa ni nujno, da pokrije dovolj dober delež problematike. Inkrementalne ocene pa imajo sposobnost prilagajanja in lahko tako bolje pokrijejo problematiko.

### 5.3 Poskusi na generatorjih z zamikom

Generatorji z zamikom nam generirajo primere neke problemske domene, ki se skozi čas (število generiranih primerov) spreminja – spreminjajo se pravila. Naš glavni cilj v tej diplomski nalogi je bil predlagati in testirati algoritme za izračun ocene zanesljivosti na prav takih problematikah. Testiranje ocen smo opravili na tri tisoč primerih. Za primerjavo s klasičnimi ocenami smo te inicializirali z uporabo prvih sto primerov. Podrobni rezultati so predstavljeni v tabelah A.17 – A.20 v dodatku A. Grafični prikaz spreminjanja korelacije skozi število primerov je prikazan v slikah A.5 – A.8. V tabeli 5.5 je prikazana statistična značilnost korelacije dejanske napake in ocene zanesljivosti.

Tabela 5.5: Odstotek značilno pozitivnih in negativnih korelacij med oceno zanesljivosti in dejansko napako napovedi.

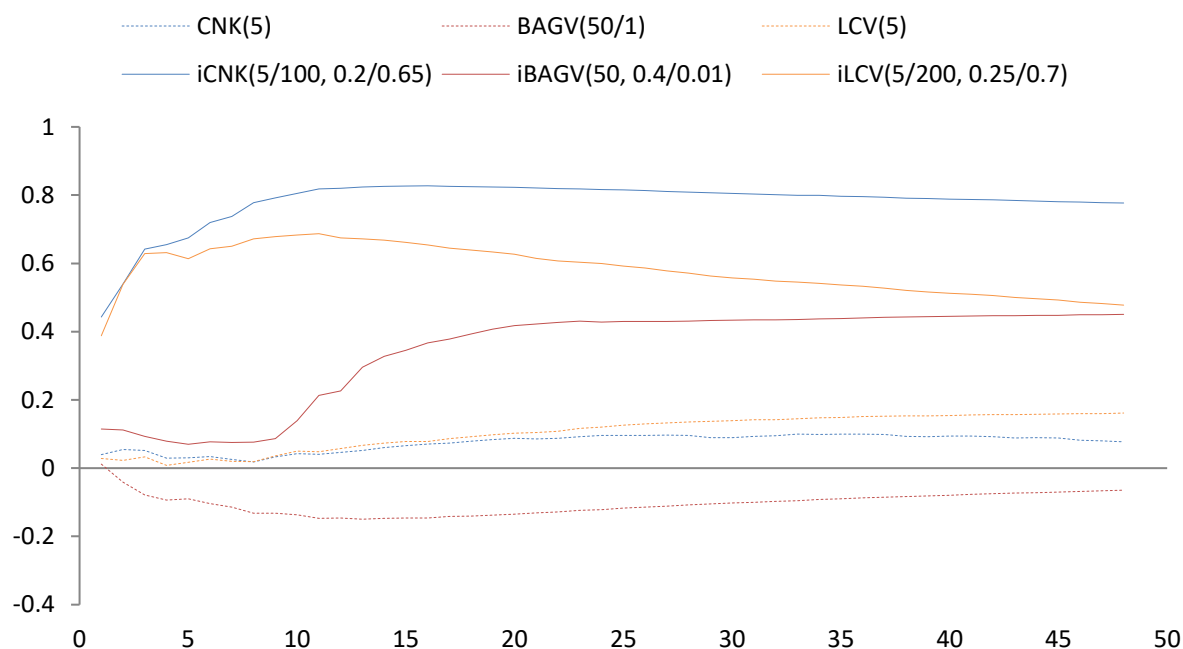
[%]	CNK +/-	BAGV +/-	LCV +/-	iCNK +/-	iBAGV +/-	iLCV +/-
iNN	88/12	62/12	88/0	100/0	88/12	100/0
HT	88/12	75/12	88/0	100/0	62/12	100/0
BAG	88/12	75/12	88/0	100/0	88/12	100/0
OBOOST	75/12	75/0	62/12	100/0	75/25	88/12
<b>povprečje</b>	<b>85/12</b>	<b>72/9</b>	<b>82/3</b>	<b>100/0</b>	<b>78/15</b>	<b>97/3</b>

Pričakovano, večina ocen korelira slabše kot pri generatorjih brez zamika. Razlog za to je spreminjanje pravil v problemski domeni. Edini oceni, ki sta izboljšali odstotek značilno pozitivnih korelacij sta iCNK in iLCV. Ocena iCNK v vseh primerih korelira značilno pozitivno in se je izkazala za najboljšo. Sledi ji ocena iLCV, ki v 97% korelira značilno pozitivno in v 3% značilno negativno. Tudi ocena iBAGV se je spremembam bolj prilagodila kot BAGV, vendar se ji je tudi povečal odstotek značilno negativnih korelacij.

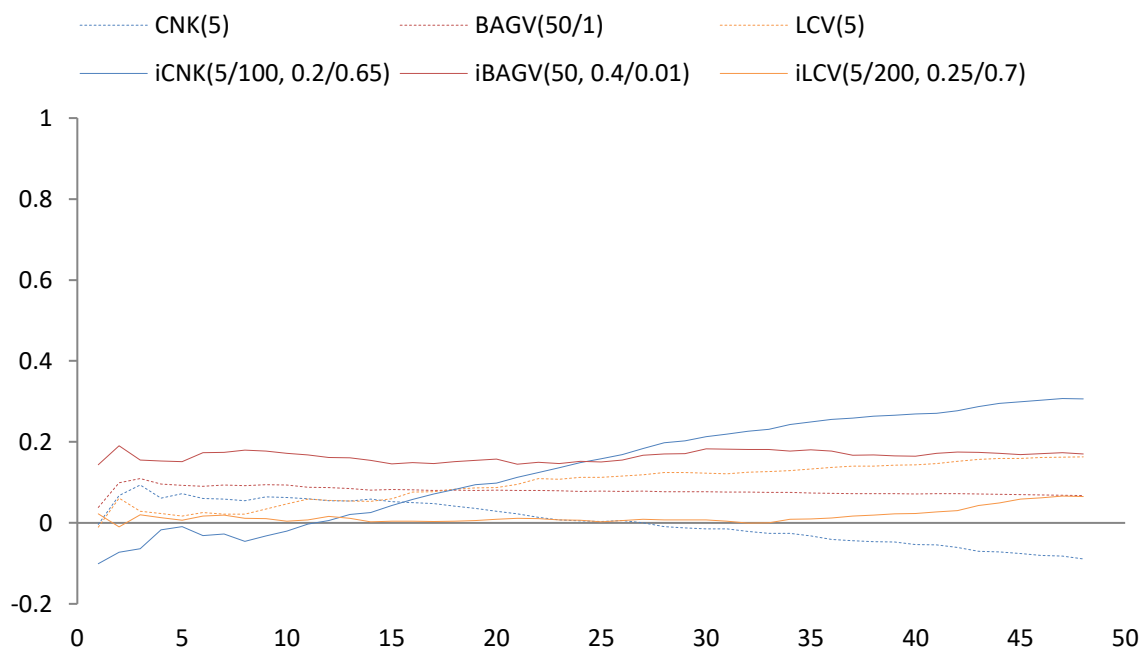
Tabela 5.6: Delež boljše ali slabše statistične značilnosti v procentih.

[%]	boljše	slabše	enako	neznačilno
iCNK	69	22	6	3
iBAGV	50	12	16	22
iLCV	19	44	22	16

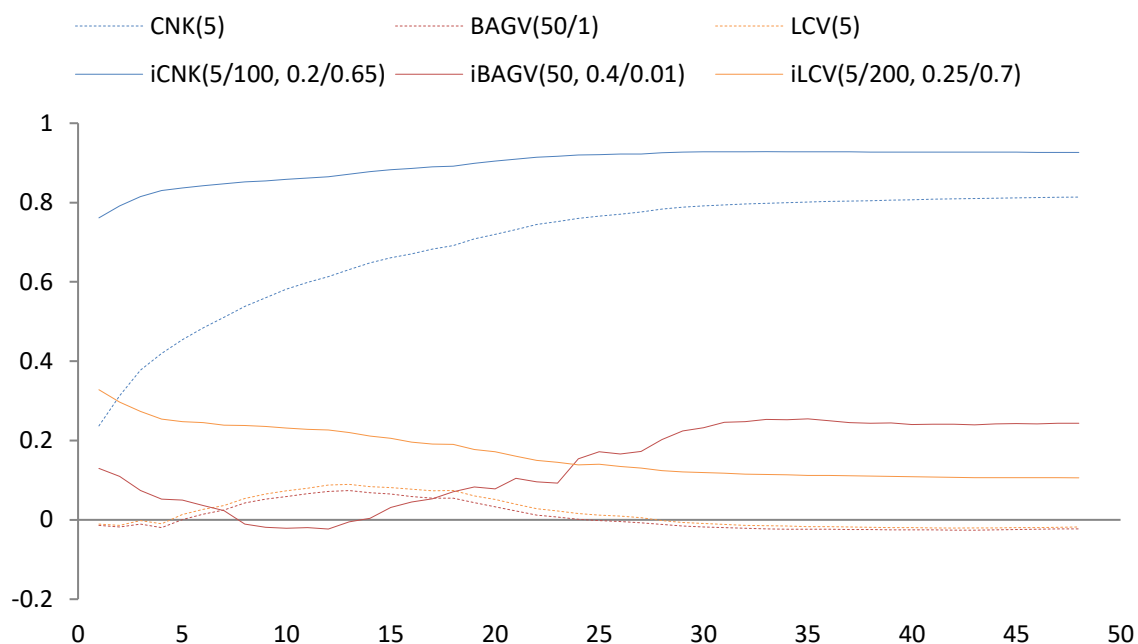
V tabeli 5.6 vidimo, da ocena iCNK v 69% primerih korelira z višjim korelacijskim koeficientom od ocene CNK, v 22% z nižjim, v 6% sta korelacijska koeficienta po z-testu enaka in v 3% vsaj eden od korelacijskih koeficientov ni statistično značilen, zato nadaljnja primerjava ni smotrna. Ocena iBAGV ima v 50% poskusov višji statistično značilen korelacijski koeficient od ocene BAGV, v 12% nižji in v 16% primerov sta po z-testu korelacijska koeficienta enaka. V 22% poskusov pa vsaj ena od ocen BAGV in iBAGV nima statistično značilnega korelacijskega koeficienta in nadaljnja primerjava ni smotrna. iLCV pa čeprav kaže boljšo statistično značilnost, kaže tudi slabšo korelacijo, saj v 44% poskusov korelira z nižjim korelacijskim koeficientom od ocene LCV. V slikah 5.4 – 5.6 je prikazanih nekaj primerov gibanja korelacije skozi primere.



Slika 5.4: Spreminjaje korelacijskega koeficienta pri generatorju STAGGER z zamikom na vsakih 60 primerov pri uporabi klasifikatorja naivni Bayes.



Slika 5.5: Spreminjaje korelacijskega koeficienta pri generatorju Agrawal z zamikom na vsakih 60 primerov pri uporabi klasifikatorja naivni Bayes.



Slika 5.6: Spreminjaje korelacijskega koeficienta pri generatorju Hyperplane z zamikom na vsakih 60 primerov pri uporabi klasifikatorja naivni Bayes.

Na problematikah z zamikom nam inkrementalne ocene ponujajo večji delež statistično značilnih pozitivnih korelacij, ki so v večini primerov tudi višje, kot korelacije klasičnih ocen. Pridobljeni rezultati nakazujejo na ustreznost algoritmov za prilagajanje spremembam.

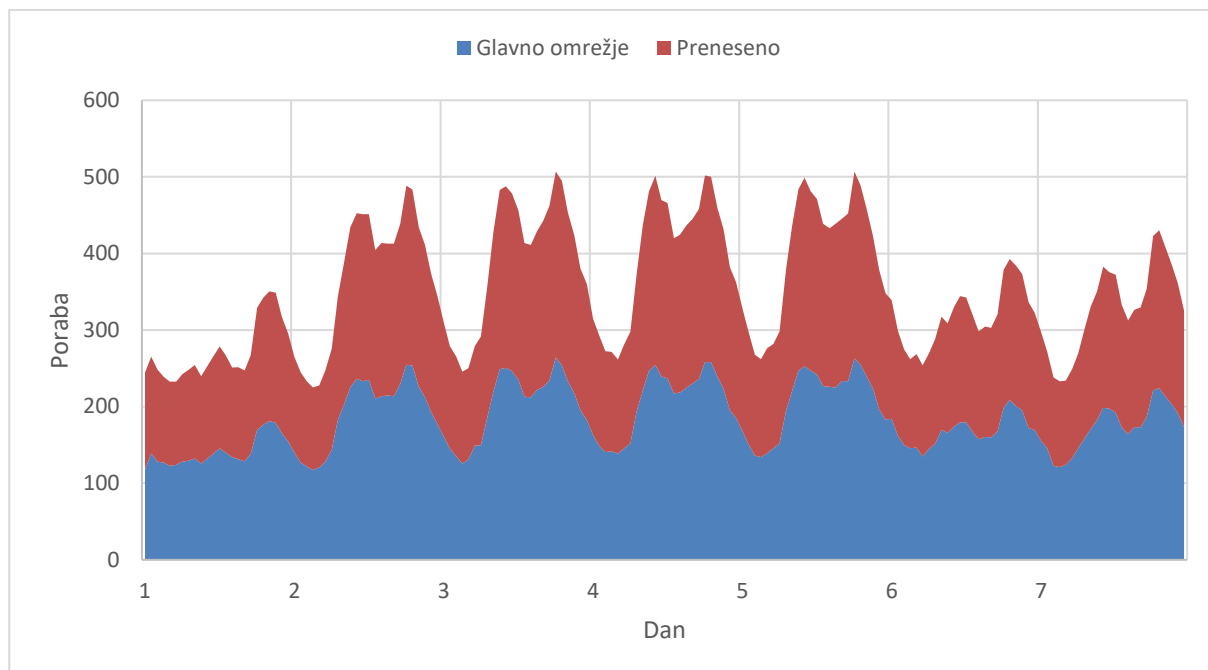


## 6 Aplikacija na realni domeni

Razvite ocene zanesljivosti smo testirali tudi na realni domeni napajanja [10] nekega italijanskega dobavitelja električne energije. V tej domeni je na vsako uro zabeležena poraba toka iz dveh virov: glavnega omrežja in pretvorjenega iz drugih omrežij. V domeni so podatki treh let od leta 1995 do 1998. Naloga strojnega učenja je, da se odloči, kateri uri dneva pripadajo določeni vrednosti porabe. Konceptualni zamik v tej domeni je v glavnem usmerjen z letnim časom, vremenom, delom dneva (jutro, večer) in razliko med delovnimi dnevi ali vikendom. V domeni je 29,928 primerov. Vsak primer ima 2 atributa in pripada enemu od 24 razredov (ure dneva).

### 6.1 Lastnosti domene

Primere v domeni opisujeta 2 decimalna številčna atributa in številčni razred, ki predstavlja uro dneva in zavzema vrednosti od 0 do 23. Slika 6.1 prikazuje razporeditev vrednosti opisnih atributov po urah dneva za sedem zaporednih dni. Opazimo lahko, da imata opisna atributa večjo vrednost od sedme pa do dvajsete ure. Prav tako sta obe vrednosti v nekaterih dneh nižji od drugih (vikend ali delovni dan). Poleg omenjenih dveh sprememb, ki so vidne v sliki 6.1, je prisotnih še več dejavnikov, ki vplivajo na vrednosti porabe (letni čas, ...). Te vse spremembe v domeni predstavljajo konceptualni zamik.



Slika 6.1: Razporeditev porabe toka iz glavnega omrežja in prenesenega iz drugih omrežij po urah za sedem zaporednih dni.

## 6.2 Testiranje ocen zanesljivosti

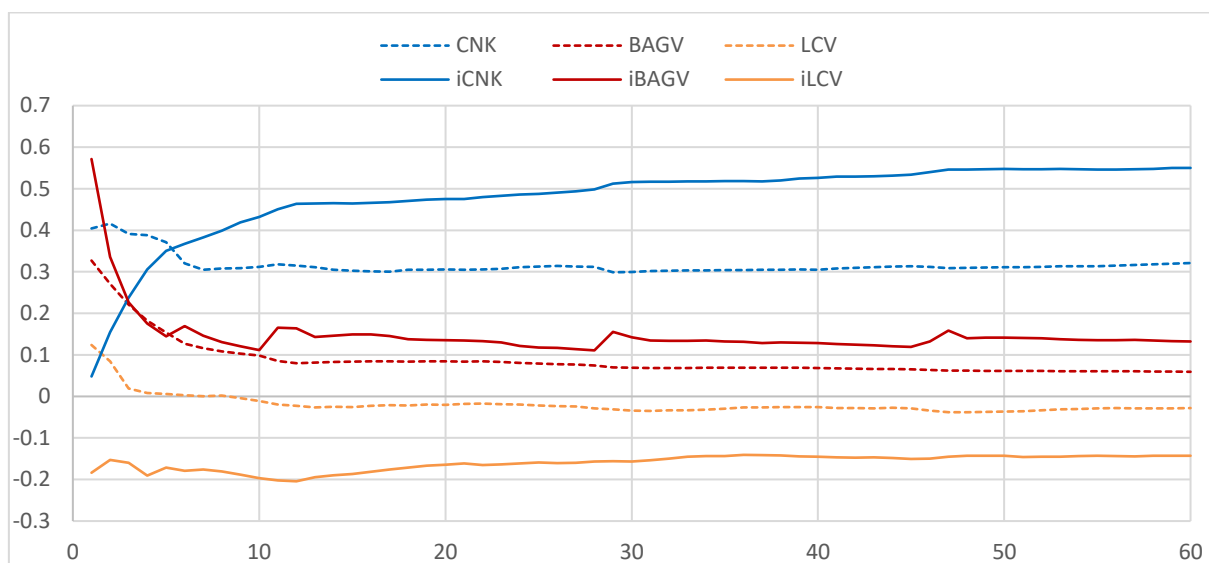
Tako inkrementalne kot klasične ocene zanesljivosti smo testirali na štirih inkrementalnih klasifikacijskih modelih na domeni napajanja. Rezultati so predstavljeni v tabeli 6.1.

Tabela 6.1: Vrednosti korelacije med dejansko napako napovedi in oceno zanesljivosti pri uporabi na realni domeni napajanja. Statistično značilno pozitivne korelacije so osenčene s svetlo sivo in negativne s temno sivo.

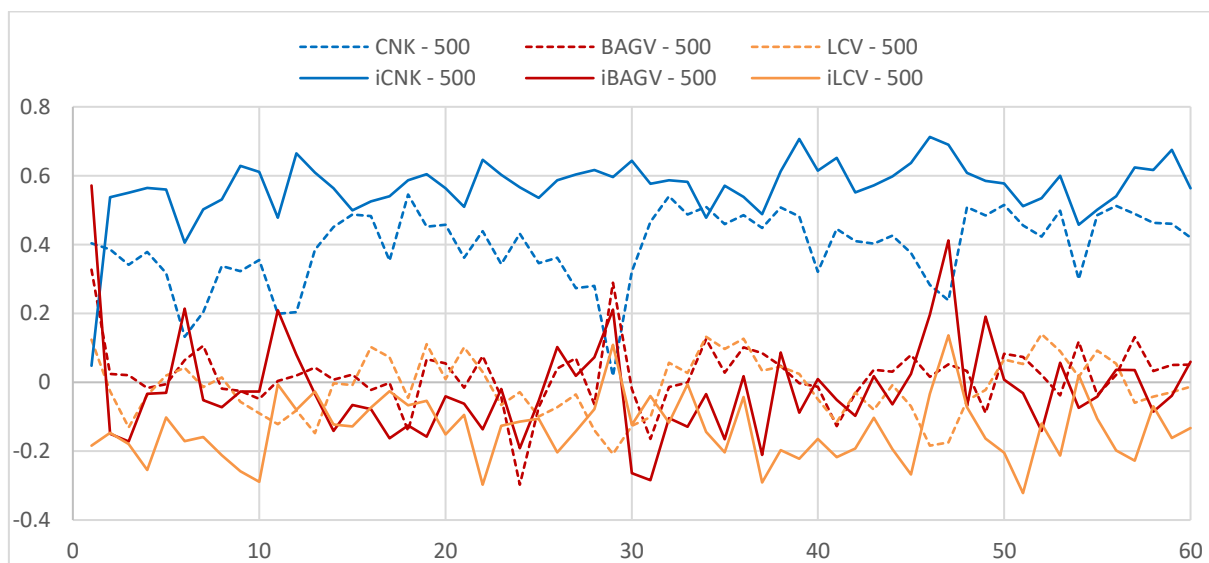
	CNK	BAGV	LCV	iCNK	iBAGV	iLCV
iNN	0.321	0.059	-0.028	0.550	0.133	-0.143
HT	0.321	0.059	-0.039	0.552	0.137	-0.143
BAG	0.328	0.045	-0.030	0.555	0.126	-0.143
OBOOST	0.123	0.005	-0.068	0.340	0.053	-0.074

Vse ocene razen LCV in iLCV imajo pozitivno statistično značilno korelacijo. Po uspešnosti si po vrstnem redu sledijo ocene iCNK, CNK, iBAGV, BAGV, LCV in iLCV. Na sliki 6.2 je grafično prikazano spreminjanje korelacije med dejansko napako in oceno zanesljivosti skozi

vse primere domene na vseh 500 primerih. Slika 6.3 pa prikazuje spreminjaje korelacije med dejansko napako in oceno zanesljivosti za zadnjih 500 primerov na vsakih 500 primerov.



Slika 6.2: Spreminjaje korelacijskega koeficienta med dejansko napako napovedi in oceno zanesljivosti za vse primere domene napajanje na vsakih 500 primerov.



Slika 6.3: Spreminjaje korelacijskega koeficienta med dejansko napako napovedi in oceno zanesljivosti za zadnjih 500 primerov domene napajanje na vsakih 500 primerov.



## 7 Diskusija in zaključki

Informacija o zanesljivosti posamezne napovedi je lahko pri odločanju kritičnega pomena. Klasične metode ocenjevanja zanesljivosti zahtevajo neko predznanje, ki mora v čim večji meri pokriti problemsko domeno. To pomeni, da klasične ocene zanesljivosti temeljijo na konsistentnosti problemske domene. V večini realnih primerov pa konsistentnost ni zagotovljena. Na primer upravljanje meritev senzorjev, katerim se skozi čas in pod vplivom drugih zunanjih dejavnikov lahko spremeni občutljivost in posledično rezultat meritev.

V okviru diplomskega dela smo razvili inkrementalne različice ocen zanesljivosti CNK, BAGV in LCV, ki temeljijo na inkrementalnem prilagajanju. Inkrementalno prilagajanje smo realizirali na dva načina. Pri ocenah iCNK in iLCV smo uporabili metodo prilagajanja prostora opisov in sicer tako, da smo z vsakim novim primerom množico opisov posodobili. Pri oceni iBAGV pa smo uporabili metodo drsečega okna. Pri tej metodi smo imeli primarno množico modelov, s katerimi smo računali oceno zanesljivosti. Z novimi primeri pa smo učili sekundarno množico modelov. Ko je del primarne množice zastaral, smo ga nadomestili s sekundarno množico in v sekundarni množici začeli graditi nove modele.

Algoritme inkrementalnih ocen smo preizkusili na 22 fiksnih množicah, 8 generatorjih in 8 generatorjih z zamikom. Algoritem ocene smo ovrednotili z izračunom Pearsonovega korelacijskega koeficienta med dejansko napako napovedi in izračunano oceno zanesljivosti, katere smo podkrepili s statističnimi metodami.

### 7.1 Pregled rezultatov

Rezultati testiranj so v skladu s pričakovanji. Na majhnih množicah primerov so se bolje izkazale klasične ocene. V teh primerih inkrementalne ocene še niso uspele dovolj dobro prilagoditi svojega omejenega prostora opisov primerov. Največjo podobnost s svojo klasično različico je v teh primerih dosegla ocena iBAGV. Pri večjih problemskih domenah so inkrementalne ocene v večini ponujale boljšo korelacijo z napako napovedi. V testiranju na problemskih domenah z zamikom pa so poleg boljše korelacije ponujale tudi večji delež značilno pozitivnih korelacij. Najslabše se je odrezala ocena iBAGV, ki je v 15% primerov korelirala značilno negativno. Vse inkrementalne različice ocen so se izkazale kot uspešne. Najvišjo uspešnost je dosegla ocena iCNK. Sledi ji ocena iLCV, kot najslabša ocena med inkrementalnimi pa se je izkazala ocena iBAGV. Povzetek rezultatov prikazuje tabela 7.1 .

Tabela 7.1: Povprečni odstotki značilno pozitivnih/negativnih korelacij med oceno zanesljivosti in dejansko napako napovedi

[%]	iCNK	iBAGV	iLCV	<b>povprečno</b>
NN	100/0	88/12	94/0	<b>94/4</b>
HT	94/0	69/12	94/0	<b>86/4</b>
OBAG	100/0	88/12	94/0	<b>94/4</b>
OBOOST	100/0	75/19	82/6	<b>86/8</b>
<b>povprečno</b>	<b>99/0</b>	<b>80/14</b>	<b>91/2</b>	<b>90/5</b>

## 7.2 Nadaljnje delo

Ideje za nadaljnje delo so:

- poskusiti uporabiti vrednost ocene zanesljivosti za popravek prvotne napovedi in tako izboljšati klasifikacijsko točnost;
- poskušati preslikati zalogo vrednosti ocen na interval  $[0, 1]$  in tako približati predstavo ocene verjetnostni obliki;
- raziskati ali je možno pri inkrementalnem prilagajanju prostora novemu primeru uporabiti kakšne mere, ki bi lahko vrednost prilagoditve samodejno uravnavale in tako omogočale hitrejšo in boljšo prilagoditev problematiki;
- razvoj pristopa za samodejno izbero najbolj primerne ocene;
- organizacija razvitih algoritmov v okolju .Net v modularno knjižnico in prispevanje paketa v javno dostopno zbirko.

## Literatura

- [1] I. Kononenko. *Strojno učenje*. Fakulteta za računalništvo in informatiko, Ljubljana (1997)
- [2] Z. Bosnić, *Ocenjevanje zanesljivosti posameznih napovedi z analizo občutljivosti regresijskih modelov*, doktorska disertacija, Fakulteta za računalništvo in informatiko, Ljubljana (2007)
- [3] Albert Bifet, Geoff Holmes, Richard Kirkby, Bernhard Pfahringer (2010); MOA: Massive Online Analysis; *Journal of Machine Learning Research* 11: 1601-160
- [4] S. Ren, Y. Lian, X. Zou, Incremental Naive Bayesian Learning Algorithm based on Classification Contribution Degree, *Journal of computers*, vol. 9, no. 8, August 2014: 1967 - 1974
- [5] D. Pevec, *Ocenjevanje zanesljivosti posameznih klasifikacij z lokalnimi metodami*, diplomsko delo, Fakulteta za računalništvo in informatiko, Ljubljana (2009)
- [6] Utgoff, P. E. (1989). Incremental induction of decision trees. *Machine Learning*, 4, 161-186.
- [7] P. Domingos, G. Hulten, Mining high-speed data streams. *In Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 71–80, Boston, MA, 2000. ACM Press
- [8] N. Oza, S. Russell, Online bagging and boosting, *In Artificial Intelligence and Statistics 2001*, 105–112, Morgan Kaufmann (2001)
- [9] A. Asuncion, D.J. Newman, *UCI Machine Learning Repository*, dostopno na naslovu <http://archive.ics.uci.edu/ml> (2007) (z dne 14. 2. 2016)
- [10] Y. Chen, E. Keogh, B. Hu, N. Begum, A. Bagnall, A. Mueen, G. Batista (2015), *The UCR Time Series Classification Archive*, dostopno na naslovu [www.cs.ucr.edu/~eamonn/time\\_series\\_data/](http://www.cs.ucr.edu/~eamonn/time_series_data/) (z dne 1. 8. 2016)





## Dodatek A

Tabela A.1: Lastnosti uporabljenih testnih domen in generatorjev.

Domena	Število primerov	Število atributov	Število razredov
flag	194	10	8
glass	214	10	6
haberman	306	4	2
heart-disease	303	14	5
hepatitis	155	20	2
iris	150	5	3
lung-cancer	32	57	3
monks-3	432	7	2
post-operative	90	9	3
sonar	208	61	2
soybean-large	307	36	19
tic-tac-toe	958	10	2
wine	178	13	3
zoo	101	17	7
generator agrawal	$\infty$	10	2
generator hyperplane	$\infty$	11	2
generator LED	$\infty$	25	10
generator random RBF	$\infty$	11	2
generator random tree	$\infty$	11	2
generator SEA	$\infty$	4	2
generator STAGGER	$\infty$	4	2
generator waveform	$\infty$	22	3

Tabela A.2: Nastavitve in inicializacija generatorjev z zamikom v MOA.

Domena	Nastavitve
generator agrawal	ConceptDriftStream -s (generators.AgrawalGenerator -f 3) -d (generators.AgrawalGenerator -f 5) -w 500 -p 500
generator hyperplane	HyperPlaneGenerator -k 2 -t 0.1
generator LED	LedGeneratorDrift -d 7
generator random RBF	RandomRBFGeneratorDrift -s 0.77 -c 5
generator random tree	ConceptDriftStream
generator SEA	ConceptDriftStream -s (generators.SEAGenerator -f 1) -d (generators.SEAGenerator -f 4)
generator STAGGER	ConceptDriftStream -s (generators.STAGGERGenerator -f 1) -d (generators.STAGGERGenerator -f 3) -w 500 -p 500
generator waveform	WaveformGeneratorDrift -d 40 -n true

Tabela A.3: Test ocene  $iCNK(5/100, \tau/0,4)$  ocene z različnimi vrednostnimi parametra  $\tau$ .

Tabela prikazuje povprečno korelacijo med zadnjimi 60 dejanskimi napakami in oceno na vsakih 60 primerov, skupno 3000 primerov. Domene, označene z zvezdico (\*), imajo zamik. Statistično značilne korelacije ( $\alpha \leq 0.05$ ) so označene s svetlo sivo barvo (značilno pozitivne korelacije) in temno sivo barvo (značilno negativne korelacije).

Generator	iCNK (0.01)	iCNK (0.1)	iCNK (0.15)	iCNK (0.2)	iCNK (0.25)
agrawal	0.030	0.024	0.046	0.030	-0.043
hyperplane	0.884	0.884	0.884	0.884	0.884
LED	0.702	0.714	0.716	0.715	0.709
random RBF	0.889	0.888	0.887	0.887	0.886
random tree	0.641	0.643	0.646	0.641	0.652
SEA	0.525	0.522	0.507	0.530	0.504
STAGGER	0.844	0.844	0.844	0.844	0.844
waveform	0.554	0.554	0.554	0.554	0.554
agrawal*	0.167	0.200	0.210	0.219	0.257
hyperplane*	0.929	0.929	0.929	0.929	0.929
LED*	0.706	0.715	0.726	0.708	0.705
rand RBF*	0.735	0.735	0.735	0.735	0.734
rand tree*	0.637	0.632	0.654	0.659	0.656
SEA*	0.519	0.534	0.523	0.532	0.513
STAGGER*	0.777	0.777	0.777	0.777	0.777
waveform*	0.429	0.429	0.429	0.429	0.429
<b>POVPREČJE</b>	<b>0.623</b>	<b>0.626</b>	<b>0.629</b>	<b>0.630</b>	<b>0.624</b>
<b>POVPREČJE*</b>	<b>0.612</b>	<b>0.619</b>	<b>0.623</b>	<b>0.623</b>	<b>0.625</b>

Tabela A.4: Test ocene  $iCNK(5/100, 0,2/\lambda)$  ocene z različnimi vrednostnimi parametra  $\lambda$ .

Tabela prikazuje povprečno korelacijo med zadnjimi 60 dejanskimi napakami in oceno na vsakih 60 primerov, skupno 3000 primerov. Domene, označene z zvezdico (\*), imajo zamik. Statistično značilne korelacije ( $\alpha \leq 0.05$ ) so označene s svetlo sivo barvo (značilno pozitivne korelacije) in temno sivo barvo (značilno negativne korelacije).

Generator	iCNK (0.1)	iCNK (0.35)	iCNK (0.5)	iCNK (0.65)	iCNK (0.8)
agrawal	-0.016	-0.036	0.039	0.063	0.009
hyperplane	0.877	0.878	0.875	0.878	0.879
LED	0.707	0.711	0.716	0.702	0.708
random RBF	0.887	0.882	0.884	0.890	0.895
random tree	0.647	0.658	0.632	0.655	0.648
SEA	0.535	0.539	0.521	0.512	0.510
STAGGER	0.844	0.844	0.844	0.844	0.844
waveform	0.539	0.566	0.568	0.574	0.595
agrawal*	0.174	0.176	0.227	0.306	0.195
hyperplane*	0.926	0.927	0.929	0.926	0.926
LED*	0.697	0.707	0.717	0.710	0.703
rand RBF*	0.739	0.736	0.744	0.747	0.755
rand tree*	0.640	0.655	0.655	0.648	0.648
SEA*	0.543	0.532	0.527	0.537	0.519
STAGGER*	0.777	0.777	0.777	0.777	0.777
waveform*	0.397	0.429	0.453	0.472	0.482
<b>POVPREČJE</b>	<b>0.620</b>	<b>0.624</b>	<b>0.632</b>	<b>0.640</b>	<b>0.631</b>
<b>POVPREČJE*</b>	<b>0.612</b>	<b>0.617</b>	<b>0.629</b>	<b>0.640</b>	<b>0.626</b>

Tabela A.5: Test ocene iBAGV(50, 0,4/ $\eta$ ) ocene z različnimi vrednostnimi parametra  $\lambda$ .

Tabela prikazuje povprečno korelacijo med zadnjimi 60 dejanskimi napakami in oceno na vsakih 60 primerov, skupno 3000 primerov. Domene, označene z zvezdico (\*), imajo zamik. Statistično značilne korelacije ( $\alpha \leq 0.05$ ) so označene s svetlo sivo barvo (značilno pozitivne korelacije) in temno sivo barvo (značilno negativne korelacije).

Generator	iBAGV (0.01)	iBAGV (0.1)	iBAGV (0.2)	iBAGV (0.3)	iBAGV (0.4)
agrawal	0.142	0.170	0.170	0.151	0.036
hyperplane	0.074	0.071	0.111	0.129	0.108
LED	0.214	0.169	0.011	-0.084	-0.148
random RBF	0.050	0.075	0.118	0.150	0.162
random tree	0.119	0.133	0.170	0.155	0.123
SEA	0.090	0.112	0.150	0.148	0.172
STAGGER	0.230	0.347	0.511	0.617	0.535
waveform	0.078	0.141	0.242	0.293	0.305
agrawal*	0.163	0.027	0.171	0.166	0.012
hyperplane*	0.036	0.078	0.212	0.246	0.272
LED*	0.216	0.174	0.017	-0.107	-0.149
rand RBF*	0.222	0.219	0.199	0.182	0.206
rand tree*	0.070	0.084	0.136	0.125	0.070
SEA*	0.027	0.054	0.084	0.108	0.106
STAGGER*	0.055	-0.142	0.159	0.458	0.471
waveform*	0.090	0.130	0.250	0.277	0.285
<b>POVPREČJE</b>	<b>0.117</b>	<b>0.115</b>	<b>0.170</b>	<b>0.188</b>	<b>0.160</b>
<b>POVPREČJE*</b>	<b>0.110</b>	<b>0.078</b>	<b>0.154</b>	<b>0.182</b>	<b>0.159</b>

Tabela A.6: Test ocene iBAGV(50,  $\lambda/0,3$ ) ocene z različnimi vrednostnimi parametra  $\eta$ .

Tabela prikazuje povprečno korelacijo med zadnjimi 60 dejanskimi napakami in oceno na vsakih 60 primerov, skupno 3000 primerov. Domene, označene z zvezdico (\*), imajo zamik. Statistično značilne korelacije ( $\alpha \leq 0.05$ ) so označene s svetlo sivo barvo (značilno pozitivne korelacije) in temno sivo barvo (značilno negativne korelacije).

Generator	iBAGV (0)	iBAGV (0.1)	iBAGV (0.2)	iBAGV (0.3)	iBAGV (0.4)
agrawal	0.103	0.068	0.135	0.139	0.114
hyperplane	0.162	0.152	0.143	0.135	0.117
LED	-0.127	-0.126	-0.112	-0.107	-0.089
random RBF	0.124	0.139	0.134	0.137	0.144
random tree	0.152	0.159	0.153	0.155	0.150
SEA	0.166	0.148	0.141	0.173	0.163
STAGGER	0.590	0.570	0.604	0.621	0.611
waveform	0.290	0.284	0.293	0.280	0.274
agrawal*	0.160	0.157	0.149	0.172	0.156
hyperplane*	0.263	0.254	0.256	0.250	0.234
LED*	-0.132	-0.123	-0.107	-0.114	-0.091
rand RBF*	0.215	0.188	0.192	0.191	0.190
rand tree*	0.106	0.113	0.118	0.129	0.122
SEA*	0.104	0.086	0.093	0.106	0.111
STAGGER*	0.450	0.438	0.454	0.446	0.456
waveform*	0.296	0.288	0.290	0.283	0.281
<b>POVPREČJE</b>	<b>0.183</b>	<b>0.175</b>	<b>0.184</b>	<b>0.187</b>	<b>0.184</b>
<b>POVPREČJE*</b>	<b>0.183</b>	<b>0.175</b>	<b>0.181</b>	<b>0.183</b>	<b>0.182</b>

Tabela A.7: Test ocene iLCV(5/100,  $\tau/0,6$ ) ocene z različnimi vrednostnimi parametra  $\tau$ .

Tabela prikazuje povprečno korelacijo med zadnjimi 60 dejanskimi napakami in oceno na vsakih 60 primerov, skupno 3000 primerov. Domene, označene z zvezdico (\*), imajo zamik. Statistično značilne korelacije ( $\alpha \leq 0.05$ ) so označene s svetlo sivo barvo (značilno pozitivne korelacije) in temno sivo barvo (značilno negativne korelacije).

Generator	iLCV (0.01)	iLCV (0.1)	iLCV (0.25)	iLCV (0.3)	iLCV (0.45)
agrawal	-0.008	-0.032	0.034	-0.005	-0.018
hyperplane	0.130	0.130	0.130	0.130	0.118
LED	0.068	0.078	0.081	0.122	0.077
random RBF	0.127	0.126	0.165	0.172	0.100
random tree	0.084	0.085	0.119	0.076	0.078
SEA	0.082	0.070	0.136	0.156	0.086
STAGGER	0.814	0.814	0.814	0.817	0.427
waveform	0.230	0.230	0.230	0.264	0.326
agrawal*	0.012	-0.016	0.037	0.013	0.113
hyperplane*	0.120	0.120	0.120	0.120	0.114
LED*	0.078	0.085	0.066	0.103	0.086
rand RBF*	0.210	0.210	0.211	0.188	0.169
rand tree*	0.077	0.065	0.099	0.092	0.096
SEA*	0.105	0.120	0.145	0.134	0.137
STAGGER*	0.478	0.478	0.478	0.477	0.436
waveform*	0.162	0.162	0.162	0.162	0.199
<b>POVPREČJE</b>	<b>0.173</b>	<b>0.170</b>	<b>0.189</b>	<b>0.189</b>	<b>0.159</b>
<b>POVPREČJE*</b>	<b>0.155</b>	<b>0.153</b>	<b>0.165</b>	<b>0.161</b>	<b>0.169</b>

Tabela A.8: Test ocene  $iLCV(5/100, 0,25/\lambda)$  ocene z različnimi vrednostnimi parametra  $\lambda$ .

Tabela prikazuje povprečno korelacijo med zadnjimi 60 dejanskimi napakami in oceno na vsakih 60 primerov, skupno 3000 primerov. Domene, označene z zvezdico (\*), imajo zamik. Statistično značilne korelacije ( $\alpha \leq 0.05$ ) so označene s svetlo sivo barvo (značilno pozitivne korelacije) in temno sivo barvo (značilno negativne korelacije).

Generator	iLCV (0.4)	iLCV (0.5)	iLCV (0.6)	iLCV (0.7)	iLCV (0.8)
agrawal	-0.015	-0.022	0.035	-0.004	-0.012
hyperplane	0.123	0.110	0.130	0.124	0.110
LED	0.097	0.057	0.096	0.081	0.076
random RBF	0.137	0.113	0.165	0.182	0.121
random tree	0.101	0.104	0.083	0.082	0.090
SEA	0.116	0.136	0.136	0.120	0.125
STAGGER	0.814	0.814	0.814	0.814	0.814
waveform	0.223	0.246	0.230	0.250	0.234
agrawal*	0.135	0.055	0.037	0.065	0.069
hyperplane*	0.112	0.100	0.120	0.106	0.109
LED*	0.071	0.077	0.095	0.098	0.071
rand RBF*	0.208	0.197	0.211	0.187	0.215
rand tree*	0.073	0.073	0.069	0.087	0.087
SEA*	0.167	0.185	0.145	0.190	0.158
STAGGER*	0.478	0.478	0.478	0.478	0.478
waveform*	0.126	0.168	0.162	0.190	0.162
<b>POVPREČJE</b>	<b>0.186</b>	<b>0.181</b>	<b>0.188</b>	<b>0.190</b>	<b>0.182</b>
<b>POVPREČJE*</b>	<b>0.171</b>	<b>0.167</b>	<b>0.165</b>	<b>0.175</b>	<b>0.169</b>



Tabela A.9: Korelacijski koeficient med oceno zanesljivosti in dejansko napako napovedi klasifikacijskega modela naivni Bayes na fiksnih množicah. Statistično značilne korelacije ( $\alpha \leq 0.05$ ) so označene s svetlo sivo barvo (značilno pozitivne korelacije) in temno sivo barvo (značilno negativne korelacije).

Množica	CNK	BAGV	LCV	iCNK	iBAGV	iLCV
generator agrawal	0.161	0.176	-0.007	0.114	0.243	-0.062
flag	0.346	0.075	0.082	0.020	0.006	-0.052
glass	0.906	0.233	0.042	0.746	0.252	0.495
haberman	0.235	0.036	0.255	0.201	0.056	0.025
heart-disease	0.479	0.339	0.285	0.313	0.389	0.038
hepatitis	0.319	0.383	0.027	0.056	0.341	0.141
generator hyperplane	0.952	0.253		0.937	0.239	0.119
iris	0.952	0.266	0.022	0.924	0.303	0.657
generator LED	0.867	-0.080	0.262	0.770	-0.054	0.192
lung-cancer	0.124	-0.263	0.124	0.350	-0.111	0.124
monks-3	0.912	0.554		0.690	0.251	0.181
post-operative	0.585	0.111	0.383	0.758	0.090	0.261
generator rand RBF	0.963	0.154	0.188	0.955	0.131	0.291
generator rand tree	0.363	0.153	0.154	0.587	0.174	0.116
generator SEA	0.338	0.236	0.254	0.505	0.228	0.262
generator STAGGER	0.914	0.619		0.777	0.547	0.823
sonar	0.982	0.117	0.088	0.876	0.277	0.047
soybean-large	0.791	0.813	0.086	0.759	0.685	-0.046
tic-tac-toe	0.892	0.330	0.430	0.755	0.078	-0.093
generator waveform	0.552	0.289	0.428	0.663	0.094	0.311
wine	0.190	0.182	0.072	0.180	0.224	0.013
zoo	0.843	0.337	0.399	0.867	0.495	0.338
<b>POVPREČJE</b>	<b>0.621</b>	<b>0.242</b>	<b>0.188</b>	<b>0.582</b>	<b>0.224</b>	<b>0.190</b>

Tabela A.10: Korelacijski koeficient med oceno zanesljivosti in dejansko napako napovedi klasifikacijskega modela Hoeffding Tree na fiksnih množicah. Statistično značilne korelacije ( $\alpha \leq 0.05$ ) so označene s svetlo sivo barvo (značilno pozitivne korelacije) in temno sivo barvo (značilno negativne korelacije).

Množica	CNK	BAGV	LCV	iCNK	iBAGV	iLCV
generator agrawal	0.161	0.126	-0.021	0.114	0.238	-0.069
flag	0.374	0.197	0.271	-0.100	0.165	0.322
glass	0.906	0.233	0.039	0.864	0.196	0.607
haberman	0.223	0.179	0.222	0.235	0.194	-0.109
heart-disease	0.479	0.325	0.293	0.255	0.352	0.025
hepatitis	0.319	0.327	0.035	0.170	0.369	0.034
generator hyperplane	0.952	0.307		0.937	0.216	0.119
iris	0.952	0.311	0.179	0.924	0.327	0.671
generator LED	0.867	-0.073	0.308	0.770	-0.045	0.177
lung-cancer	0.101	-0.246	0.053	0.425	-0.159	0.053
monks-3	0.905	0.335	0.395	0.717	0.276	0.214
post-operative	0.754	-0.030	0.452	0.855	0.050	0.448
generator rand RBF	0.963	0.118	0.060	0.955	0.137	0.303
generator rand tree	0.363	0.084	0.205	0.701	0.163	0.008
generator SEA	0.338	0.225	0.214	0.505	0.195	0.197
generator STAGGER	0.914	0.592		0.777	0.509	0.761
sonar	0.982	0.141	-0.004	0.867	0.076	0.232
soybean-large	0.794	0.811	0.074	0.771	-0.174	-0.064
tic-tac-toe	0.887	0.275	0.403	0.756	0.132	-0.069
generator waveform	0.552	0.277	0.389	0.663	0.132	0.291
wine	0.190	0.158	0.113	0.132	0.236	0.086
zoo	0.844	0.469	0.387	0.865	0.651	0.366
<b>POVPREČJE</b>	<b>0.628</b>	<b>0.234</b>	<b>0.203</b>	<b>0.598</b>	<b>0.193</b>	<b>0.209</b>

Tabela A.11: Korelacijski koeficient med oceno zanesljivosti in dejansko napako napovedi klasifikacijskega modela Oza Bag na fiksnih množicah. Statistično značilne korelacije ( $\alpha \leq 0.05$ ) so označene s svetlo sivo barvo (značilno pozitivne korelacije) in temno sivo barvo (značilno negativne korelacije).

Množica	CNK	BAGV	LCV	iCNK	iBAGV	iLCV
generator agrawal	0.172	0.325	-0.009	0.073	0.390	-0.086
flag	0.395	0.211	0.122	0.083	0.156	-0.046
glass	0.909	0.155	0.035	0.771	0.184	0.366
haberman	0.230	0.128	0.228	0.263	0.045	0.019
heart-disease	0.473	0.451	0.289	0.213	0.463	-0.057
hepatitis	0.273	0.382	0.048	0.207	0.329	-0.153
generator hyperplane	0.954	0.363		0.940	0.202	0.112
iris	0.946	0.307	0.031	0.890	0.302	0.665
generator LED	0.860	-0.050	0.449	0.786	-0.010	0.316
lung-cancer	0.173	-0.205	0.054	0.452	-0.021	0.054
monks-3	0.904	0.496	0.399	0.707	0.282	0.158
post-operative	0.626	0.216	0.479	0.735	0.172	0.426
generator rand RBF	0.964	0.254	0.215	0.960	0.170	0.228
generator rand tree	0.344	0.193	0.149	0.675	0.133	0.129
generator SEA	0.371	0.262	0.256	0.519	0.222	0.254
generator STAGGER	0.932	0.670		0.825	0.628	0.854
sonar	0.973	0.129	0.041	0.885	0.146	0.020
soybean-large	0.509	0.258	0.261	0.472	0.000	0.109
tic-tac-toe	0.893	0.343	0.469	0.762	0.089	-0.151
generator waveform	0.546	0.326	0.394	0.661	0.138	0.293
wine	0.199	0.382	0.065	0.187	0.359	0.160
zoo	0.841	0.374	0.462	0.857	0.483	0.390
<b>POVPREČJE</b>	<b>0.613</b>	<b>0.271</b>	<b>0.222</b>	<b>0.587</b>	<b>0.221</b>	<b>0.184</b>

Tabela A.12: Korelacijski koeficient med oceno zanesljivosti in dejansko napako napovedi klasifikacijskega modela Oza Boost na fiksnih množicah. Statistično značilne korelacije ( $\alpha \leq 0.05$ ) so označene s svetlo sivo barvo (značilno pozitivne korelacije) in temno sivo barvo (značilno negativne korelacije).

Množica	CNK	BAGV	LCV	iCNK	iBAGV	iLCV
generator agrawal	0.267	0.478	-0.041	0.122	0.489	-0.040
flag	0.222	0.058	0.171	0.031	0.000	-0.080
glass	0.866	0.342	-0.198	0.848	0.359	-0.227
haberman	0.320	0.116	0.069	0.353	0.060	-0.065
heart-disease	0.541	0.224	0.233	0.190	0.419	-0.105
hepatitis	0.283	0.418	-0.036	0.285	0.342	0.059
generator hyperplane	0.969	-0.021	-0.279	0.965	0.102	-0.249
iris	0.908	0.525	-0.187	0.875	0.514	-0.077
generator LED	0.873	0.091	0.414	0.778	0.036	0.275
lung-cancer	0.089	0.088	0.138	0.341	0.033	0.138
monks-3	0.955	0.694	0.029	0.825	0.712	0.033
post-operative	0.552	0.291	-0.001	0.639	0.181	0.039
generator rand RBF	0.985	0.474	0.145	0.974	0.372	0.131
generator rand tree	0.454	0.170	0.111	0.764	0.175	0.221
generator SEA	0.529	-0.104	0.242	0.572	-0.129	0.321
generator STAGGER	0.839	0.297	-0.548	0.671	0.305	-0.413
sonar	0.955	0.298	0.080	0.820	0.320	-0.014
soybean-large	0.623	0.265	0.252	0.562	-0.029	0.134
tic-tac-toe	0.949	0.148	-0.242	0.737	0.084	-0.060
generator waveform	0.588	0.400	0.250	0.627	0.419	0.233
wine	0.192	0.493	0.088	0.223	0.691	0.089
zoo	0.813	0.293	0.119	0.825	0.600	0.175
<b>POVPREČJE</b>	<b>0.626</b>	<b>0.274</b>	<b>0.037</b>	<b>0.592</b>	<b>0.275</b>	<b>0.023</b>

Tabela A.13: Korelacijski koeficient med oceno zanesljivosti in dejansko napako napovedi klasifikacijskega modela naivni Bayes na generatorjih brez zamika. Statistično značilne korelacije ( $\alpha \leq 0.05$ ) so označene s svetlo sivo barvo (značilno pozitivne korelacije) in temno sivo barvo (značilno negativne korelacije).

Množica	CNK	BAGV	LCV	iCNK	iBAGV	iLCV
generator agrawal	0.221	0.050	0.035	0.063	0.109	-0.004
generator hyperplane	0.761	0.120	0.121	0.878	0.127	0.124
generator LED	0.771	0.130	0.161	0.699	-0.100	0.087
generator rand RBF	0.829	0.056	0.109	0.890	0.134	0.182
generator rand tree	0.305	0.076	0.169	0.642	0.157	0.086
generator SEA	0.201	0.036	0.175	0.512	0.161	0.120
generator STAGGER	0.569	0.152	0.283	0.844	0.580	0.814
generator waveform	0.395	0.103	0.287	0.574	0.279	0.250
<b>POVPREČJE</b>	<b>0.506</b>	<b>0.090</b>	<b>0.168</b>	<b>0.638</b>	<b>0.181</b>	<b>0.207</b>

Tabela A.14: Korelacijski koeficient med oceno zanesljivosti in dejansko napako napovedi klasifikacijskega modela Hoeffding Tree na generatorjih brez zamika. Statistično značilne korelacije ( $\alpha \leq 0.05$ ) so označene s svetlo sivo barvo (značilno pozitivne korelacije) in temno sivo barvo (značilno negativne korelacije).

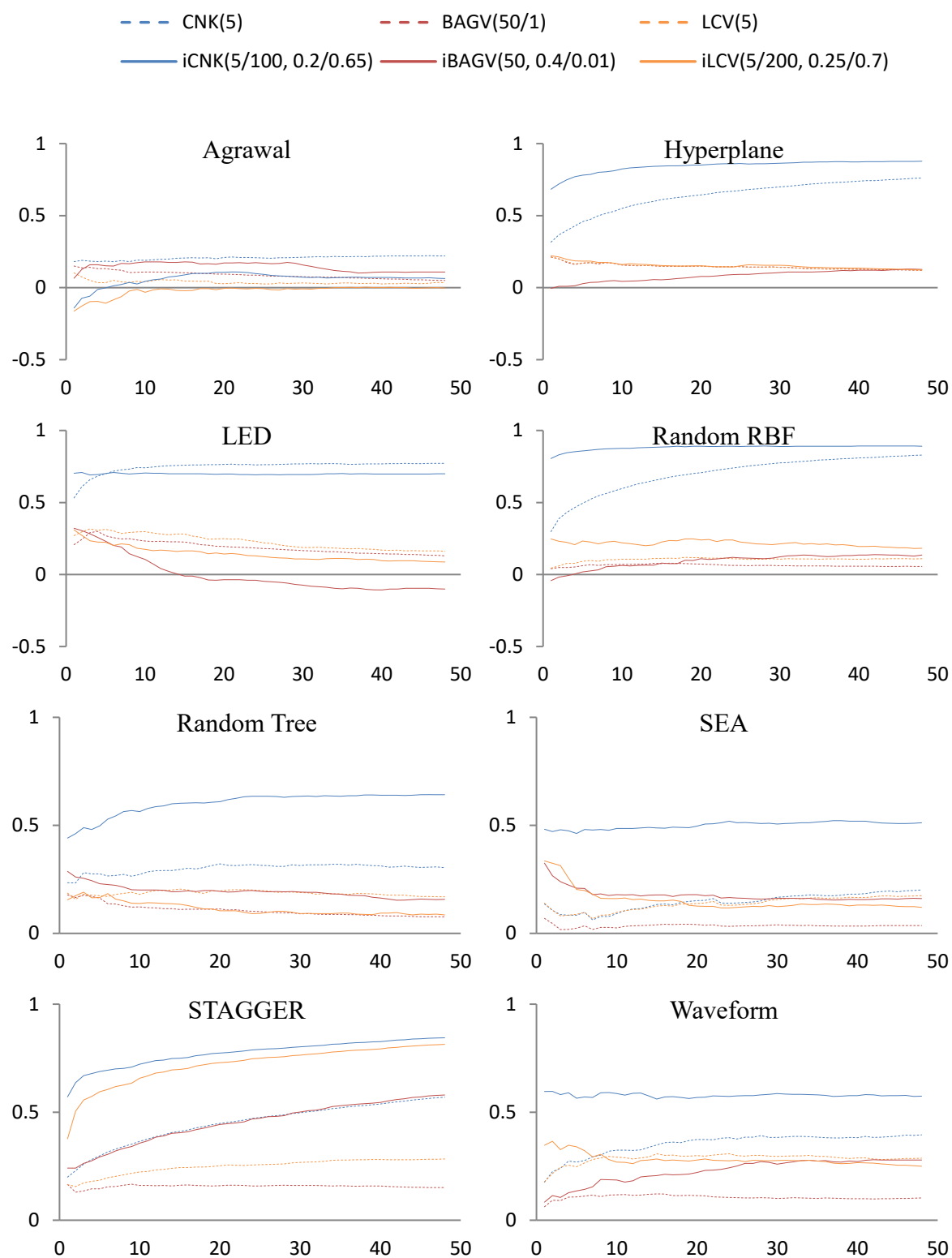
Množica	CNK	BAGV	LCV	iCNK	iBAGV	iLCV
generator agrawal	0.171	0.088	0.006	0.028	0.185	-0.001
generator hyperplane	0.761	0.120	0.121	0.878	0.123	0.142
generator LED	0.771	0.130	0.192	0.714	-0.105	0.092
generator rand RBF	0.830	0.045	0.113	0.897	0.112	0.160
generator rand tree	0.263	0.087	0.142	0.659	-0.007	0.082
generator SEA	0.203	0.035	0.176	0.518	0.117	0.127
generator STAGGER	0.569	0.190	0.283	0.844	0.652	0.811
generator waveform	0.395	0.098	0.280	0.574	0.292	0.243
<b>POVPREČJE</b>	<b>0.495</b>	<b>0.099</b>	<b>0.164</b>	<b>0.639</b>	<b>0.171</b>	<b>0.207</b>

Tabela A.15: Korelacijski koeficient med oceno zanesljivosti in dejansko napako napovedi klasifikacijskega modela Oza Bag na generatorjih brez zamika. Statistično značilne korelacije ( $\alpha \leq 0.05$ ) so označene s svetlo sivo barvo (značilno pozitivne korelacije) in temno sivo barvo (značilno negativne korelacije).

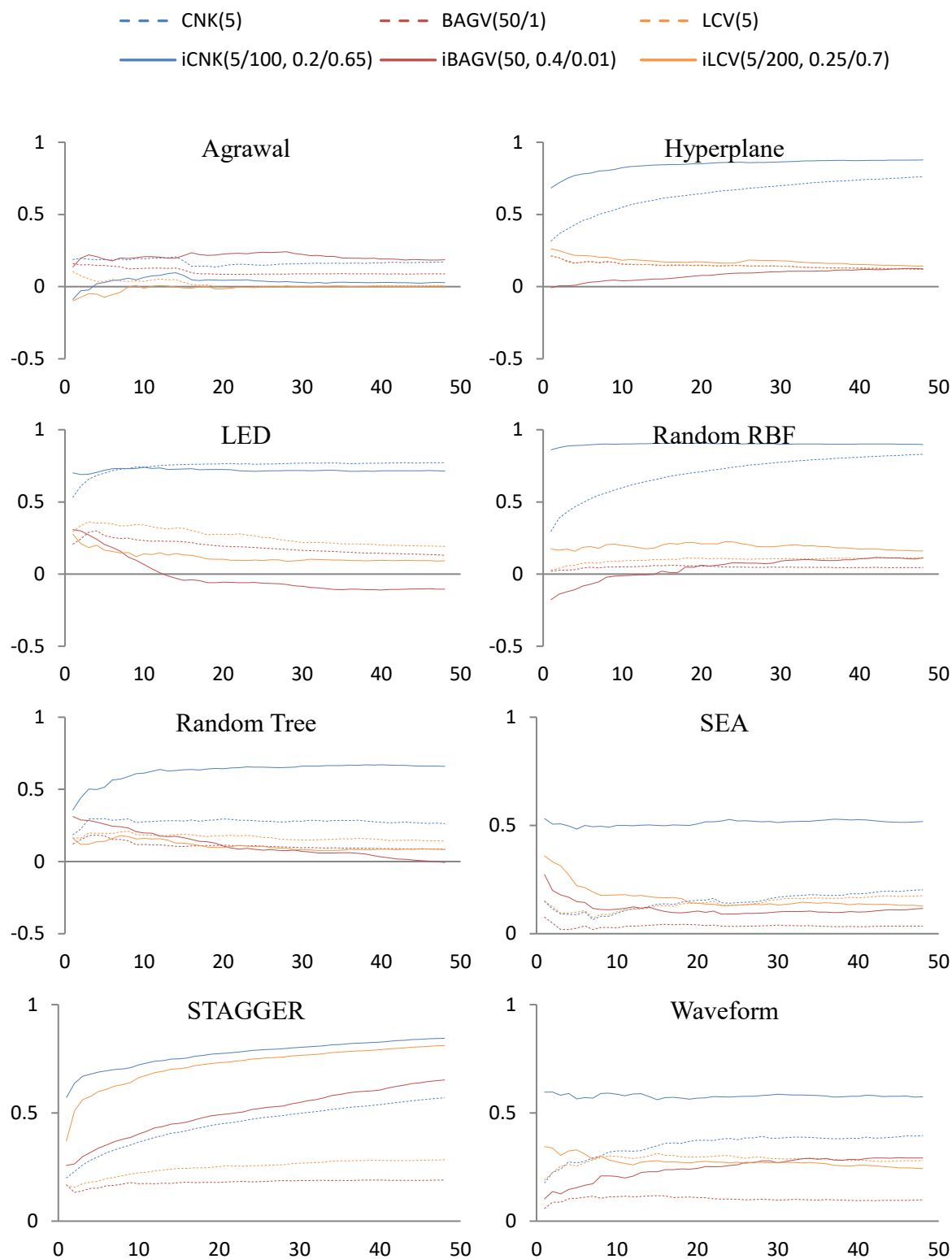
Množica	CNK	BAGV	LCV	iCNK	iBAGV	iLCV
generator agrawal	0.235	0.110	0.044	0.067	0.144	0.004
generator hyperplane	0.780	0.137	0.138	0.883	0.219	0.152
generator LED	0.775	0.122	0.250	0.713	-0.180	0.154
generator rand RBF	0.835	0.051	0.102	0.897	0.165	0.178
generator rand tree	0.300	0.062	0.170	0.649	0.119	0.038
generator SEA	0.200	0.022	0.174	0.517	0.140	0.129
generator STAGGER	0.585	0.154	0.273	0.854	0.714	0.823
generator waveform	0.392	0.078	0.288	0.572	0.343	0.258
<b>POVPREČJE</b>	<b>0.513</b>	<b>0.092</b>	<b>0.180</b>	<b>0.644</b>	<b>0.208</b>	<b>0.217</b>

Tabela A.16: Korelacijski koeficient med oceno zanesljivosti in dejansko napako napovedi klasifikacijskega modela Oza Boost na generatorjih brez zamika. Statistično značilne korelacije ( $\alpha \leq 0.05$ ) so označene s svetlo sivo barvo (značilno pozitivne korelacije) in temno sivo barvo (značilno negativne korelacije).

Množica	CNK	BAGV	LCV	iCNK	iBAGV	iLCV
generator agrawal	0.370	0.209	0.081	0.128	0.587	-0.008
generator hyperplane	0.649	0.189	0.087	0.879	0.279	0.035
generator LED	0.785	0.151	0.262	0.730	-0.192	0.220
generator rand RBF	0.526	0.038	0.029	0.861	0.136	0.041
generator rand tree	0.251	0.065	0.181	0.628	0.048	0.109
generator SEA	0.244	0.003	0.177	0.526	0.033	0.168
generator STAGGER	0.597	0.451	-0.072	0.625	0.166	0.433
generator waveform	0.451	0.096	0.229	0.556	0.193	0.237
<b>POVPREČJE</b>	<b>0.484</b>	<b>0.150</b>	<b>0.122</b>	<b>0.617</b>	<b>0.156</b>	<b>0.154</b>

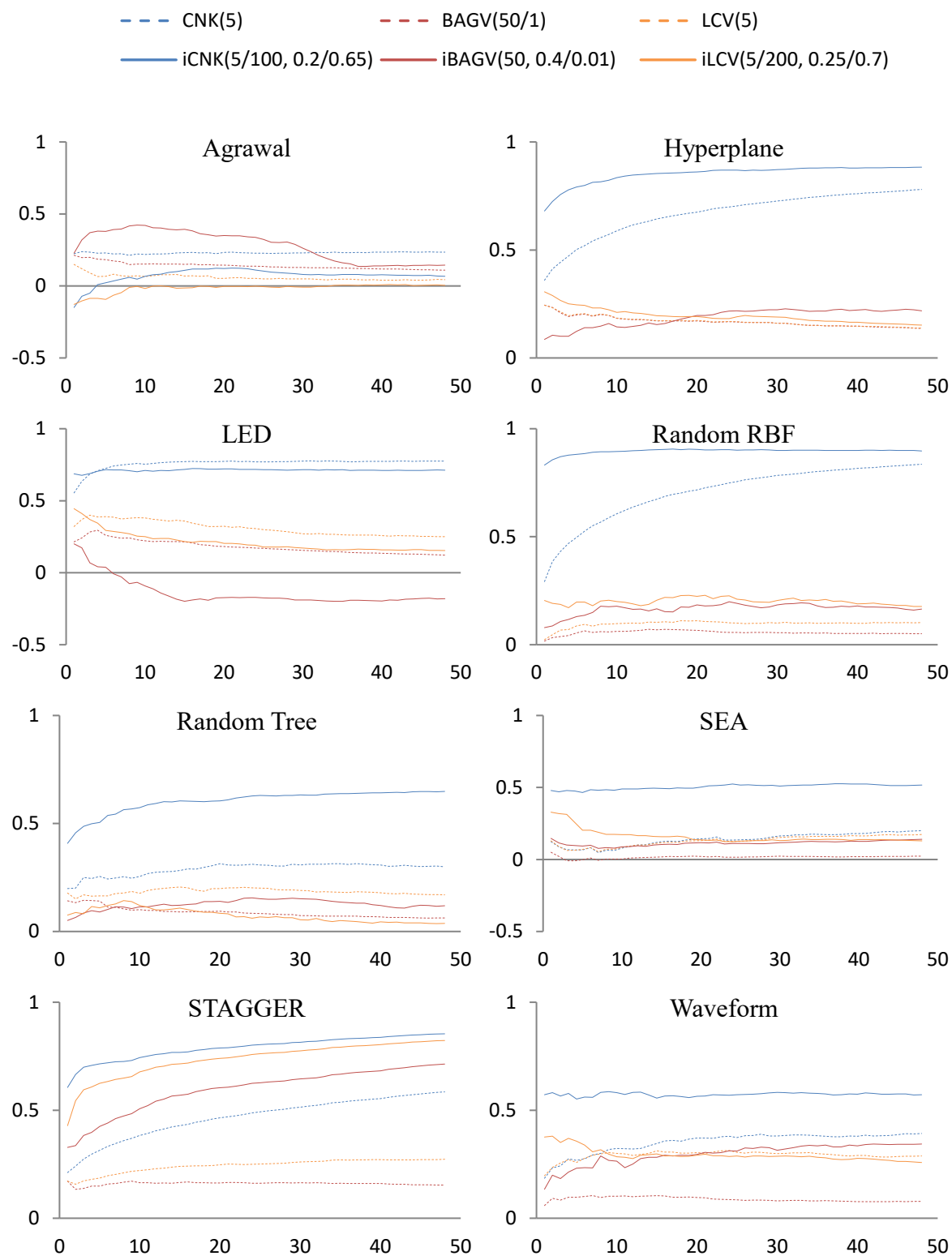


Slika A.1: Spreminjanje korelacije med dejansko napako in oceno zanesljivosti na vsakih 60 primerov pri generatorjih brez zamikom in inkrementalnem klasifikatorjem naivni Bayes.

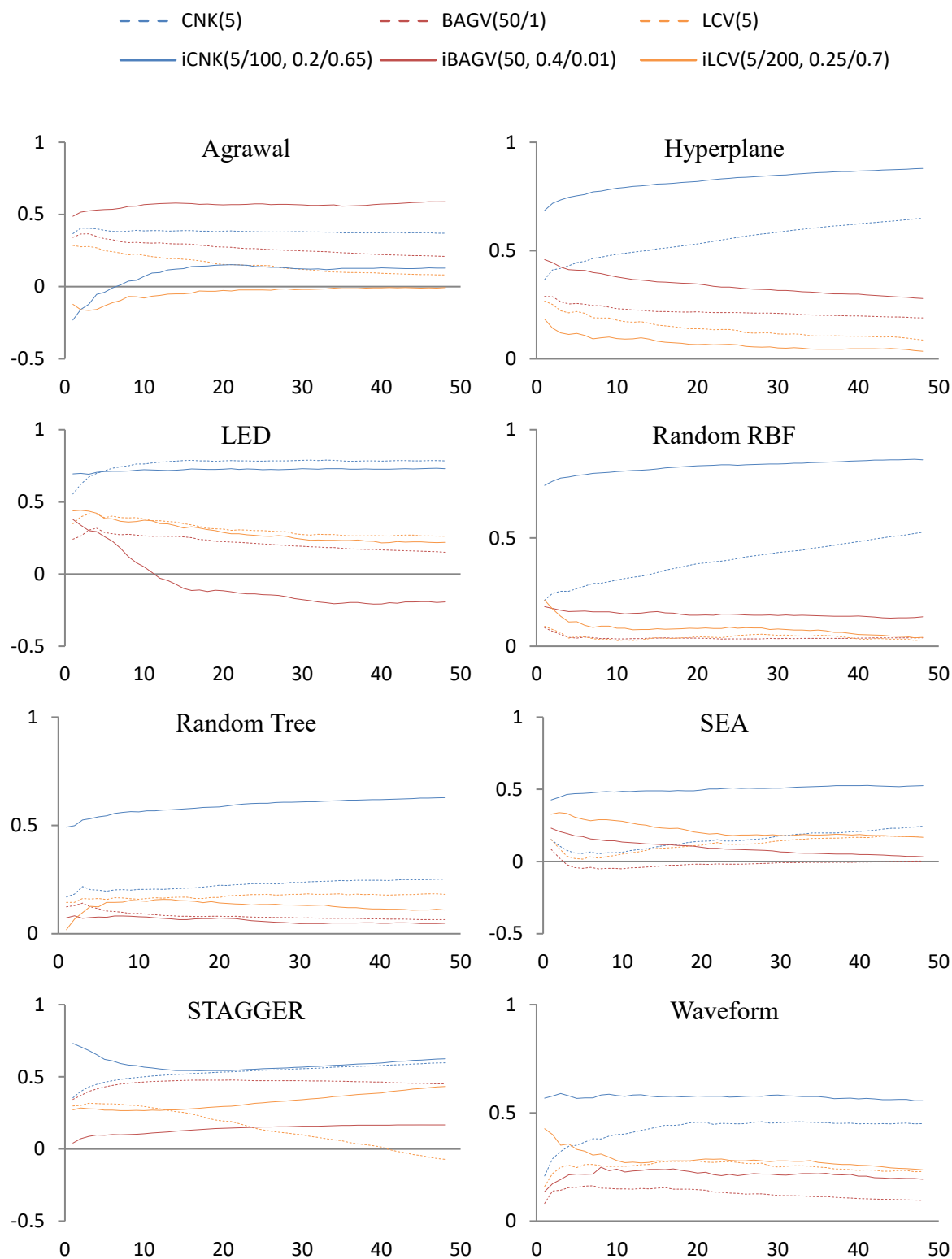


Slika A.2: Spreminjanje korelacije med dejansko napako in oceno zanesljivosti na vsakih 60 primerov pri generatorjih brez zamikom in inkrementalnem klasifikatorjem Hoeffding Tree.





Slika A.3: Spreminjanje korelacije med dejansko napako in oceno zanesljivosti na vsakih 60 primerov pri generatorjih brez zamikom in inkrementalnem klasifikatorjem Oza Bag.



Slika A.4: Spreminjanje korelacije med dejansko napako in oceno zanesljivosti na vsakih 60 primerov pri generatorjih brez zamikom in inkrementalnem klasifikatorjem Oza Boost.

Tabela A.17: Korelacijski koeficient med oceno zanesljivosti in dejansko napako napovedi klasifikacijskega modela naivni Bayes na generatorjih z zamikom. Statistično značilne korelacije ( $\alpha \leq 0.05$ ) so označene s svetlo sivo barvo (značilno pozitivne korelacije) in temno sivo barvo (značilno negativne korelacije).

Množica	CNK	BAGV	LCV	iCNK	iBAGV	iLCV
generator agrawal	-0.089	0.067	0.163	0.306	0.171	0.065
generator hyperplane	0.814	-0.022	-0.018	0.926	0.243	0.106
generator LED	0.771	0.131	0.161	0.705	-0.107	0.072
generator rand RBF	0.790	0.200	0.357	0.747	0.178	0.187
generator rand tree	0.240	0.059	0.072	0.655	0.122	0.087
generator SEA	0.228	0.035	0.200	0.537	0.104	0.190
generator STAGGER	0.077	-0.064	0.161	0.777	0.451	0.478
generator waveform	0.422	0.135	0.285	0.472	0.277	0.190
<b>POVPREČJE</b>	<b>0.407</b>	<b>0.068</b>	<b>0.173</b>	<b>0.641</b>	<b>0.180</b>	<b>0.172</b>

Tabela A.18: Korelacijski koeficient med oceno zanesljivosti in dejansko napako napovedi klasifikacijskega modela Hoeffding Tree na generatorjih z zamikom. Statistično značilne korelacije ( $\alpha \leq 0.05$ ) so označene s svetlo sivo barvo (značilno pozitivne korelacije) in temno sivo barvo (značilno negativne korelacije).

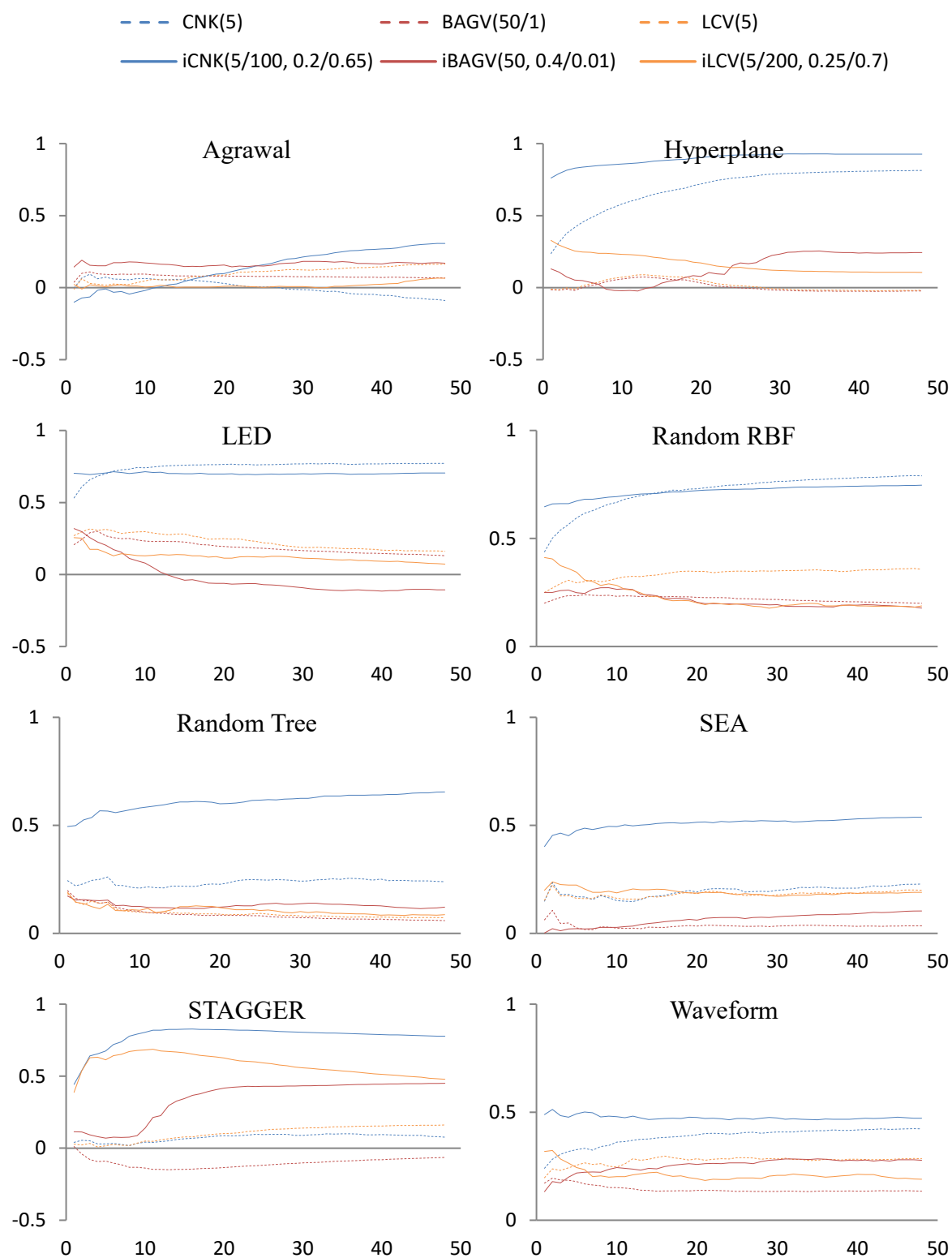
Množica	CNK	BAGV	LCV	iCNK	iBAGV	iLCV
generator agrawal	-0.089	0.068	0.183	0.306	0.147	0.068
generator hyperplane	0.845	0.009	0.011	0.924	-0.003	0.120
generator LED	0.771	0.130	0.192	0.711	-0.103	0.142
generator rand RBF	0.753	0.154	0.485	0.683	0.251	0.332
generator rand tree	0.226	0.056	0.078	0.629	-0.009	0.087
generator SEA	0.230	0.037	0.210	0.541	0.137	0.188
generator STAGGER	0.077	-0.065	0.164	0.777	0.271	0.505
generator waveform	0.418	0.116	0.266	0.474	0.275	0.173
<b>POVPREČJE</b>	<b>0.404</b>	<b>0.063</b>	<b>0.198</b>	<b>0.631</b>	<b>0.121</b>	<b>0.202</b>

Tabela A.19: Korelacijski koeficient med oceno zanesljivosti in dejansko napako napovedi klasifikacijskega modela Oza Bag na generatorjih z zamikom. Statistično značilne korelacije ( $\alpha \leq 0.05$ ) so označene s svetlo sivo barvo (značilno pozitivne korelacije) in temno sivo barvo (značilno negativne korelacije).

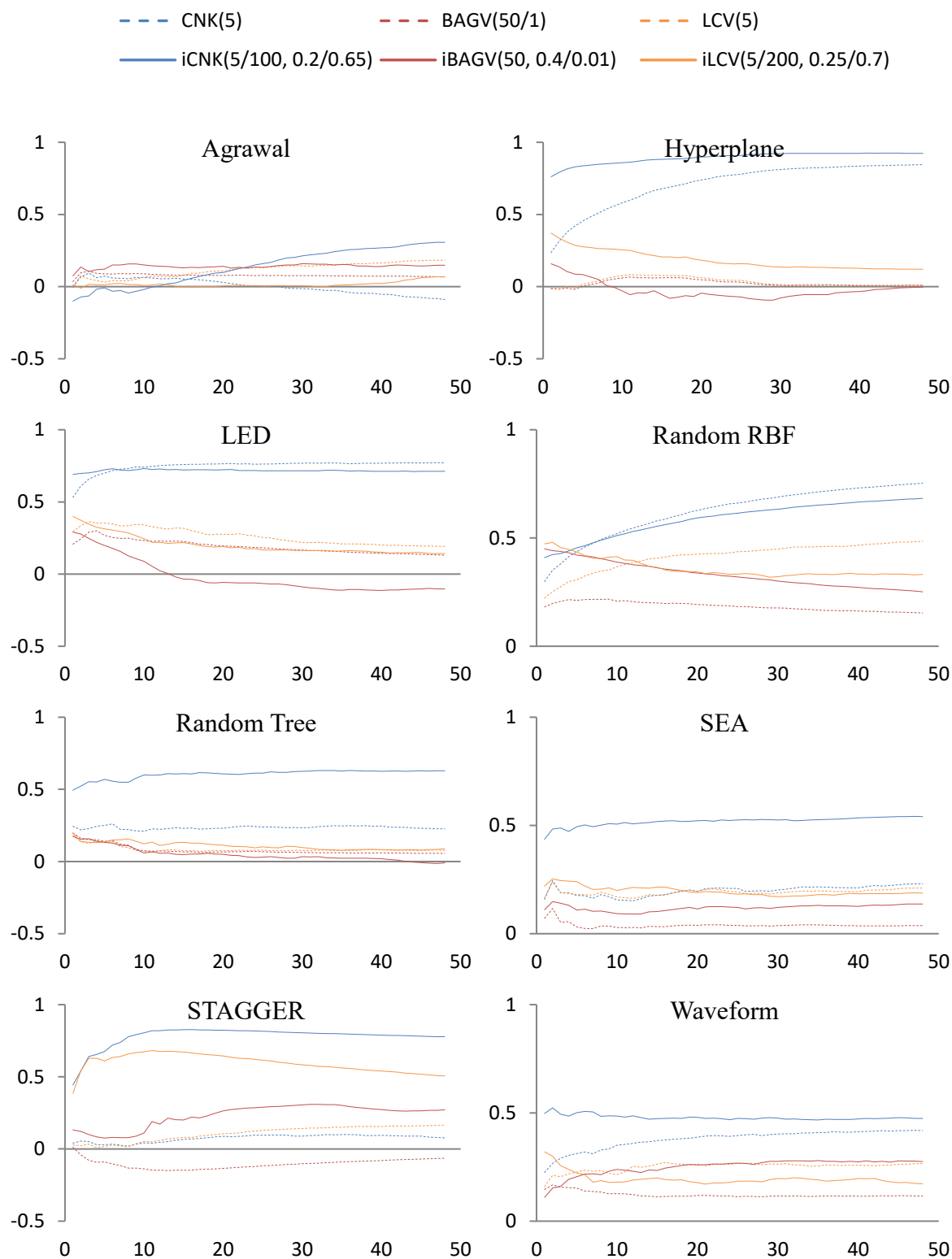
Množica	CNK	BAGV	LCV	iCNK	iBAGV	iLCV
generator agrawal	-0.086	0.073	0.171	0.326	0.096	0.073
generator hyperplane	0.822	-0.026	-0.021	0.930	0.246	0.092
generator LED	0.775	0.122	0.250	0.716	-0.197	0.186
generator rand RBF	0.773	0.113	0.315	0.764	0.083	0.191
generator rand tree	0.237	0.043	0.069	0.655	0.125	0.080
generator SEA	0.228	0.036	0.206	0.539	0.138	0.193
generator STAGGER	0.083	-0.063	0.157	0.778	0.473	0.521
generator waveform	0.424	0.105	0.290	0.470	0.332	0.192
<b>POVPREČJE</b>	<b>0.407</b>	<b>0.051</b>	<b>0.180</b>	<b>0.647</b>	<b>0.162</b>	<b>0.191</b>

Tabela A.20: Korelacijski koeficient med oceno zanesljivosti in dejansko napako napovedi klasifikacijskega modela Oza Boost na generatorjih z zamikom. Statistično značilne korelacije ( $\alpha \leq 0.05$ ) so označene s svetlo sivo barvo (značilno pozitivne korelacije) in temno sivo barvo (značilno negativne korelacije).

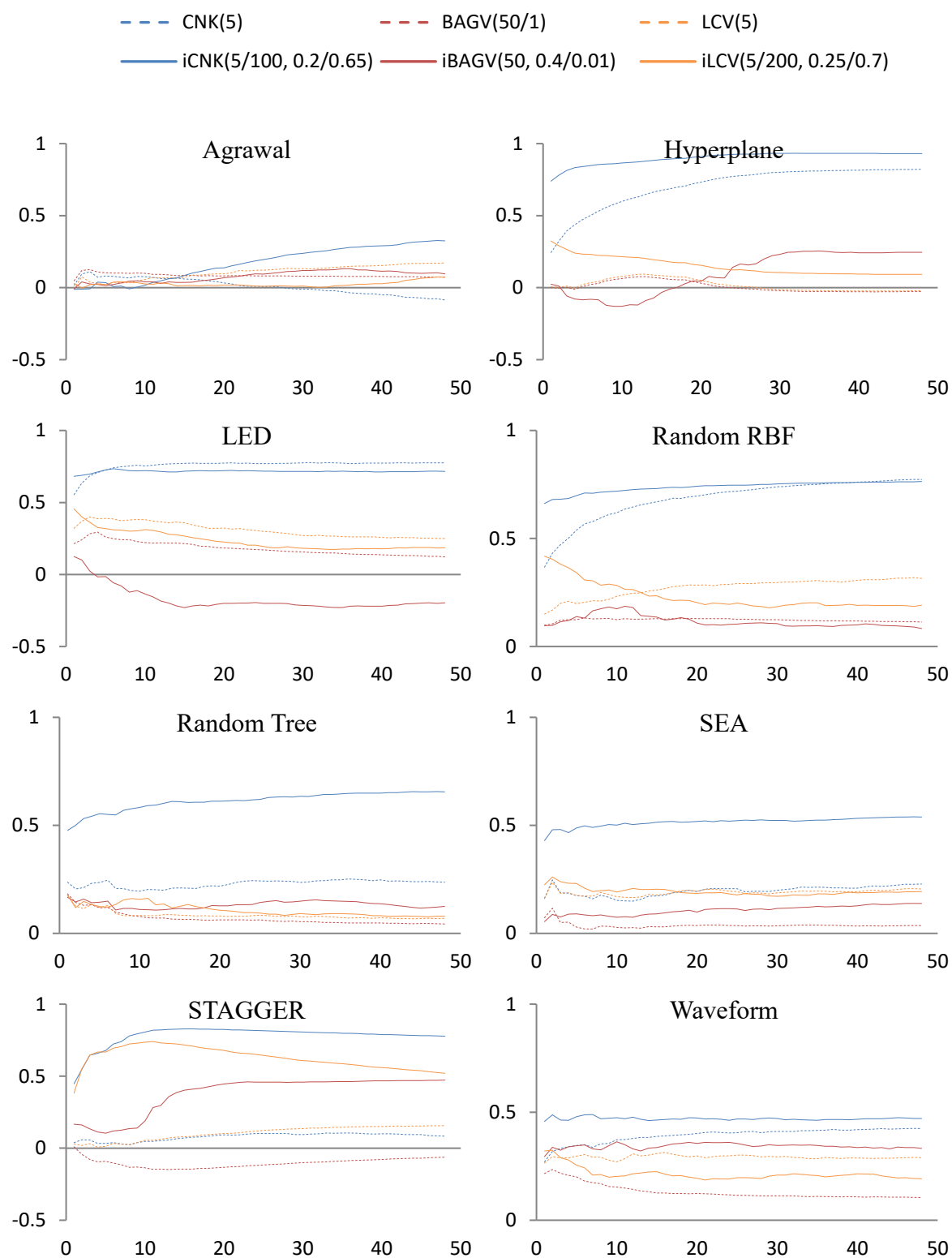
Množica	CNK	BAGV	LCV	iCNK	iBAGV	iLCV
generator agrawal	-0.110	0.085	0.087	0.386	0.140	0.106
generator hyperplane	0.583	0.094	-0.021	0.879	0.153	-0.041
generator LED	0.785	0.151	0.262	0.727	-0.198	0.205
generator rand RBF	0.573	0.028	0.021	0.509	-0.064	0.082
generator rand tree	0.233	0.050	0.065	0.613	0.051	0.145
generator SEA	0.280	0.039	0.039	0.509	0.041	0.211
generator STAGGER	0.009	-0.002	-0.097	0.747	0.402	0.112
generator waveform	0.436	0.113	0.190	0.463	0.163	0.136
<b>POVPREČJE</b>	<b>0.349</b>	<b>0.070</b>	<b>0.068</b>	<b>0.604</b>	<b>0.086</b>	<b>0.119</b>



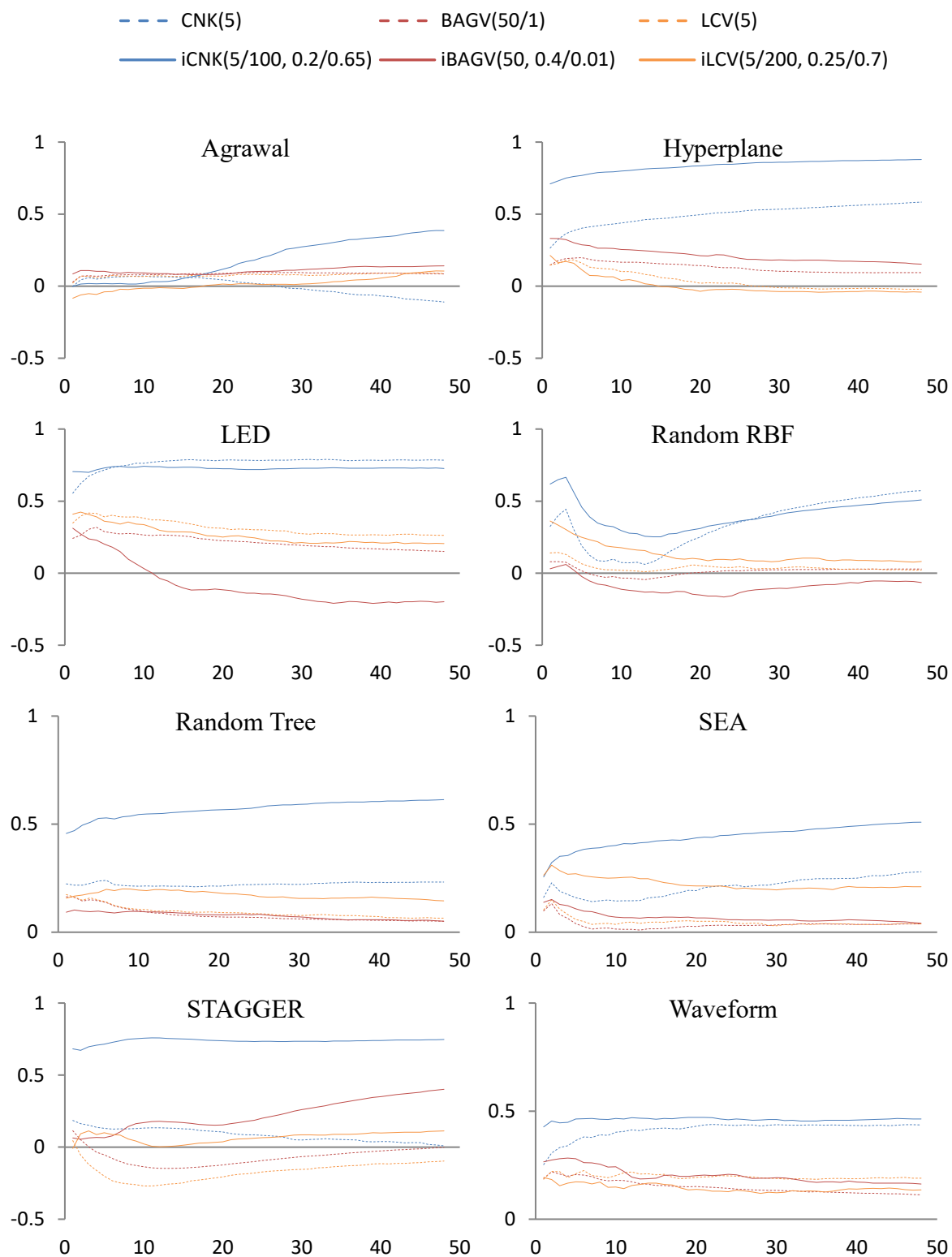
Slika A.5: Spreminjanje korelacije med dejansko napako in oceno zanesljivosti na vsakih 60 primerov pri generatorjih z zamikom in inkrementalnem klasifikatorjem naivni Bayes.



Slika A.6: Spreminjanje korelacije med dejansko napako in oceno zanesljivosti na vsakih 60 primerov pri generatorjih z zamikom in inkrementalnem klasifikatorjem Hoeffding Tree.



Slika A.7: Spreminjanje korelacije med dejansko napako in oceno zanesljivosti na vsakih 60 primerov pri generatorjih z zamikom in inkrementalnem klasifikatorjem Oza Bag.



Slika A.8: Spreminjanje korelacije med dejansko napako in oceno zanesljivosti na vsakih 60 primerov pri generatorjih z zamikom in inkrementalnem klasifikatorjem Oza Boost